

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 27, 2004, 17:04:25 ; Search time 15.1307 Seconds
(without alignments)
1593.347 Million cell updates/sec

Title: US-10-037-860-13

Perfect score: 2423

Sequence: 1 MPTLLQDWCGRGHEHLNTRC.....VESGNGNWAWDKSHPKSKAK 463

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	836.5	34.5	364	1 PMA2 HUMAN	Q9u142 homo sapien
2	826.5	34.1	364	1 PMA2 MACFA	Q9gm33 macaca fasc
3	822	33.9	353	1 PMA1 MOUSE	Q8clc8 mus musculu
4	822	33.9	353	1 PMA1 RAT	Q8vht4 rattus norv
5	818	33.8	353	1 PMA1 HUMAN	Q8nd90 homo sapien
6	768.5	31.7	351	1 MOP1 HUMAN	Q96by2 homo sapien
7	761.5	31.4	351	1 MOP1 MACFA	Q95ki4 macaca fasc
8	725.5	29.9	365	1 PMA2 MOUSE	Q8bhk0 mus musculu
9	711	29.3	352	1 MOP1 MOUSE	Q8erb6 mus musculu
10	114.5	4.7	474	1 LAM3 MOUSE	P48680 mus musculu
11	113	4.7	476	1 ENGA RHINE	Q92uk6 rhizobium m
12	113	4.7	3674	1 SPCR HUMAN	Q9nrc6 homo sapien
13	110	4.5	373	1 CK13 HUMAN	Q02833 homo sapien
14	107.5	4.4	817	1 PERQ HUMAN	O75420 homo sapien
15	106.5	4.4	301	1 SSB DEIRA	O9r751 deinococcus
16	106.5	4.4	857	1 MEDB GIALA	Q08014 giardia lam
17	105.5	4.4	813	1 TTKA DROME	P42282 drosophila
18	105	4.3	783	1 RIN1 HUMAN	Q13671 homo sapien
19	105	4.3	804	1 SYM ORYSA	Q9ztst oryza sativ
20	105	4.3	4625	1 DVI4 CHLRE	Q9amh3 chlamydomon
21	104.5	4.3	864	1 RA50 SULSO	Q97wh0 sulfolobus
22	104	4.3	483	1 ENGA BRUME	Q8yfh2 brucella me
23	104	4.3	483	1 ENGA BRUSU	Q9g2e8 brucella su
24	104	4.3	2564	1 SPCQ HUMAN	Q9h254 homo sapien
25	103	4.3	860	1 IF2 RHIL0	Q98bi8 rhizobium l
26	103	4.3	1011	1 PERQ MOUSE	Q99mr1 mus musculu
27	102.5	4.2	4367	1 DYHC NEUCR	P45443 neurospora
28	101.5	4.2	1733	1 VNDA PRVKA	P33485 pseudorabie
29	100	4.1	384	1 AIPL HUMAN	Q9nzn9 homo sapien
30	100	4.1	384	1 AIPL PANPA	Q95mn9 pan paniscu
31	99.5	4.1	600	1 LAM2 CHICK	P14732 gallus gall
32	99.5	4.1	1050	1 BUIB HUMAN	O60566 homo sapien
33	99	4.1	398	1 MPK2_CHICK	Q90891 gallus gall

RESULT 1

PM2 HUMAN
ID PM2 HUMAN STANDARD; PRT; 364 AA.
AC Q9UL42; Q94959; Q95145; Q9UL43;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Paraneoplastic antigen Ma2 (Onconeural antigen MA2) (Paraneoplastic
neural antigen MA2) (40 kDa neuronal protein).
GN PMA2 OR MA2 OR KIA0883.
OS Homo sapiens (Human).
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RP TISSUE=Cerebellum;
RC MEDLINE=99270611; PubMed=10362822;
RA Voltz R., Gultekin S.H., Rosenfeld M.R., Gerstner E., Eichen J.,
Posner J.B., Dalmay J.;
RT "A serologic marker of paraneoplastic limbic and brain-stem
encephalitis in patients with testicular cancer.";
RL New Engl. J. Med. 340:1788-1795(1999).
[2]
RN SEQUENCE FROM N.A.
RP TISSUE=Brain;
RC MEDLINE=99156230; PubMed=10048485;
RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirose M.,
Miyaajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XII.
The complete sequences of 100 new cDNA clones from brain which code
for large proteins in vitro.";
RL DNA Res. 5:355-364(1998).
[3]
RN IDENTIFICATION, AND SUBCELLULAR LOCATION.
RP MEDLINE=99158179; PubMed=10050892;
RA Dalmay J., Gultekin S.H., Voltz R., Hoard R., DesChamps T.,
Balmaceda C., Batchelor T., Gerstner E., Eichen J., Frenner J.,
Posner J.B., Rosenfeld M.R.;
RT "Ma1, a novel neuron- and testis-specific protein, is recognized by
the serum of patients with paraneoplastic neurological disorders.";
RL Brain 122:27-39(1999).
CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar.
CC -!- TISSUE SPECIFICITY: Brain specific. In some patients suffering
from cancers, it is also specifically expressed by the testicular
tumor cells.
CC -!- MISCELLANEOUS: Antibodies against PMA2 are present in sera from
patients suffering of paraneoplastic neurological disorders.
CC -!- SIMILARITY: Belongs to the PMA family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

34 98.5 4.1 692 1 MYS PODCA Q05000 podocoryne
35 98.5 4.1 1197 1 CING_HUMAN Q9p2m7 homo sapien
36 98.5 4.1 3680 1 DMD_CANFA Q97592 canis famli
37 98 4.0 1268 1 YRD6 CAEEL Q95755 caenorhabdi
38 97.5 4.0 476 1 ENGA_AGR5 Q8ud28 agrobacteri
39 97.5 4.0 888 1 M3KC_RAT Q63796 rattus norv
40 97.5 4.0 1178 1 DSRA_MOUSE Q99mu3 mus musculu
41 97 4.0 524 1 EXP9_STRPN P35599 streptococc
42 97 4.0 780 1 CUL5_RABIT Q29425 cryptocolag
43 96.5 4.0 441 1 GAG_CAEVC P33458 caprine art
44 96.5 4.0 570 1 SYG_ARCFU Q29346 archaeglob
45 96.5 4.0 714 1 CLPB_MYCGE P47597 mycoplasma

ALIGNMENTS

```

CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AF037365; AD02098.1; -.
CC EMBL; AF083114; AAF05625.1; -.
CC EMBL; AF083115; AAF05626.1; -.
CC EMBL; AF286487; AAG28165.1; -.
CC EMBL; AB020690; BAA74906.1; -.
CC Genew; HGNC:9159; PNMA2.
CC MIM; 603970; -.
CC Antigen; Tumor antigen; Nuclear protein.
CC FT DOMAIN 333 338 POLY-GLU.
CC FT CONFLICT 128 129 GV -> AL (IN REF. 1; AAD02098).
CC FT CONFLICT 141 141 LV -> I (IN REF. 1; AAD02098).
CC FT CONFLICT 257 257 T -> P (IN REF. 1; AAF05625).
CC FT CONFLICT 278 278 R -> K (IN REF. 1; AAF05626).
CC SEQUENCE 364 AA; 41509 MW; 6E417AD96E3F0E93 CRC64;

Query Match 34.5%; Score 836.5; DB 1; Length 364;
Best Local Similarity 49.7%; Pred. No. 2.4e-53;
Matches 174; Conservative 63; Mismatches 108; Indels 5; Gaps 3;

QY 1 MPTLLQDWCRCGHLNTRCMLILGIPEDCGEDEFETLQEAACHLGRYRVIGRMFRREE 60
DB 1 MALALLEDCRIMSVDQKSLMVTGIPVDYEEAEIQEVLTQKSLGRLGKIFRKQE 60
QY 61 NQAAILLEAODIDYALLPREIPKGGPWEIVKPNSDGEFLNRLNLEERRTVSDM 120
DB 61 NANAVLLELDDTDVSAIPSEVQGGVWVFKTPNQDTEFLERLNLFLKEGGQTVSGM 120
QY 121 NRVLGSDTNCSPRVTPISPEFWT--WAQTILGAVQPLLEOMLYRELRFVSGNTISIPGAL 178
DB 121 FRALQEGVSPATVPICISPELLAHLGQAMAHAPQLL-PMRYKLRVFGSAVPAPEE 179
QY 179 AFDAMLEHTTEMLQMWQVPEGEKRRRLMECLRGPALQVVGSLPASNASITVECLAAQQ 238
DB 180 SPEVWLEQATEIVKVPVTEAEKKRWLAESLRGPDALDLMHIVQADNPSISVECLFAFKQ 239
QY 239 VFGPVESHKIAQVLCCKAYOEAGEKVSFVLRLEPLLQRAVENNVSRNVTNRLKRVL 298
DB 240 VFGSLESRTAQVRYLYKTYQEEGKVSAYVLRLETLRRRAVEKRAIPRIADQVRLEQVM 299
QY 299 SGATLPDKLRDKLKMQRKPPGFALVKLLREEEWEATLGPDRRESLE 348
DB 300 AGATLNQMLWCLRELKQDGGPPSFLELMKVIREEEEAESP--ENESIE 347

RESULT 2
PMA2 MACFA
ID PMA2 MACFA STANDARD; PRT; 364 AA.
AC Q9GMU3;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DE Paraneoplastic antigen Mac2 homolog.
GN PNMA2.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OC NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT Isolation of full-length cDNA clones from macaque brain cDNA
RT libraries.
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar (By similarity).
CC -!- SIMILARITY: Belongs to the PNMA family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AB047632; BAB12156.1; -.
CC Nuclear protein.
CC FT DOMAIN 333 338 POLY-GLU.
CC FT SEQUENCE 364 AA; 41350 MW; 0CF72210D7EC1524 CRC64;

Query Match 34.1%; Score 826.5; DB 1; Length 364;
Best Local Similarity 49.1%; Pred. No. 1.3e-52;
Matches 172; Conservative 63; Mismatches 110; Indels 5; Gaps 3;

QY 1 MPTLLQDWCRCGHLNTRCMLILGIPEDCGEDEFETLQEAACHLGRYRVIGRMFRREE 60
DB 1 MALALLEDCRIMSVDQKSLMVTGIPVDYEEAEIQEVLTQKSLGRLGKIFRKQE 60
QY 61 NQAAILLEAODIDYALLPREIPKGGPWEIVKPNSDGEFLNRLNLEERRTVSDM 120
DB 61 NANAVLLELDDTDVSAIPSEVQGGVWVFKTPNQDTEFLERLNLFLKEGGQTVSGM 120
QY 121 NRVLGSDTNCSPRVTPISPEFWT--WAQTILGAVQPLLEOMLYRELRFVSGNTISIPGAL 178
DB 121 FRALQEGVSPATVPICISPELLAHLGQAMAHAPQLL-PMRYKLRVFGSAVPAPEE 179
QY 179 AFDAMLEHTTEMLQMWQVPEGEKRRRLMECLRGPALQVVGSLPASNASITVECLAAQQ 238
DB 180 SPEVWLEQATEIVKVPVTEAEKKRWLAESLRGPDALDLMHIVQADNPSISVECLFAFKQ 239
QY 239 VFGPVESHKIAQVLCCKAYOEAGEKVSFVLRLEPLLQRAVENNVSRNVTNRLKRVL 298
DB 240 VFGSLESRTAQVRYLYKTYQEEGKVSAYVLRLETLRRRAVEKRAIPRIADQVRLEQVM 299
QY 299 SGATLPDKLRDKLKMQRKPPGFALVKLLREEEWEATLGPDRRESLE 348
DB 300 AGATLNQMLWCLRELKQDGGPPSFLELMKVIREEEEAESP--ENESIE 347

RESULT 3
PMA1 MOUSE
ID PMA1 MOUSE STANDARD; PRT; 353 AA.
AC Q8CJG8; Q9CVP2;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Paraneoplastic antigen Mal homolog.
GN PNMA1 OR MA1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Brain, and Embryo;
RX MEDLINE=22354683; PubMed=12466851;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusica V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempke C.A., Setou M., Shimada K.,

```

```

RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavalon M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:563-573 (2002).
CC -1- SUBCELLULAR LOCATION: Nuclear; nucleolar (By similarity).
CC -1- SIMILARITY: Belongs to the PNMA family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AK017476; BAB30762.1; -.
DR EMBL; AK028331; BAC25885.1; -.
DR MGD; MGI:2180564; Pnma1.
KW Nuclear protein.
FT DOMAIN 336 341 POLY-GLU.
FT CONFLICT 136 136 T -> M (IN REF. 1; BAB30762).
FT SEQUENCE 353 AA; 39688 MW; 18CEDC3AC4E70939 CRC64;
SQ
Query Match 33.9%; Score 822; DB 1; Length 353;
Best Local Similarity 49.4%; Pred. No. 2.6e-52;
Matches 175; Conservative 54; Mismatches 117; Indels 8; Gaps 4;
Qy 1 MPTLLQDWCGRGHEHLNTRCMLILGIPEDCGEDFEETLQEAACHLGRYRIGRMFRREE 60
Db 1 MAMTLLDWCGRGMDVNSQALLVWGIPVNCDETIETLQAMPQVS-YRVLGRMFREE 59
Qy 61 NQAAILLELAQDIDYALLPREIPGKGGPWEIVKPRNSDGEFLNRLNRFLEERRTVSDM 120
Db 60 NAKAALLELTGVDYSLIPREMPGKGLKVKVPPTSDAFLERHLFLAREGWTQDV 119
Qy 121 NRVLGSDNTCSAPRTVISPEFTWQAOTLGAAVQLPQLQMLYRELRFVSGNTISIPGALAF 180
Db 120 ARVLGFQNPAPAPGPETPAEMLN--TLDNVIQPLVESIWYKLTFLSGKDIPGGEETF 177
Qy 181 DAWLEHTTEMLQWQVEGEGKRRLEMCRLCPALQVVSGLRASNASITVEECLAQOVF 240
Db 178 DSWLEHSENEIEEWQVSDIEKRRRLMSLGRPADVIRILKTNNAITTAELKALEQVF 237
Qy 241 GPVESHKIAQVKLCAYQEAQEKVSSFFVLRLEPLLRQAVENNVVSRNNVQTRLKRVLSG 300
Db 238 GSVESSRDAQVRFLNTYQNPGEKLSYVIRLEPLQLQKVDKGVIDKONVQARLEQVIAG 297
Qy 301 ATLPDKLRDULKMKQRKRP-PGFLALVKLLRE-----EEWEATIGPDRESLEG 349
Db 298 ANHSGALRRQLWLAGAEGPAPNLFQLLVQIREBEAKKEEAEAAALQLGLEG 351
RESULT 4
PMAL RAT STANDARD; PRT; 353 AA.
ID PMAL RAT
AC Q8VHZ4;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DE 15-MAR-2004 (Rel. 43, Last annotation update)
DE Paraneoplastic antigen Mal homolog.
GN PNMA1 OR MAL.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

```

```

OX NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
MEDLINE=99158179; PubMed=10050892;
RX Balmau J., Gultekin S.H., Voltz R., Hoard R., DesChamps T.,
RA Balmau C., Bachelot T., Gerstner E., Richen J., Frennier J.,
RA Posner J.B., Rosenfeld M.R.;
RT "Mal, a novel neuron- and testis-specific protein, is recognized by
RT the serum of patients with paraneoplastic neurological disorders.";
RL Brain 122:27-39 (1999).
CC -1- SUBCELLULAR LOCATION: Nuclear; nucleolar (By similarity).
CC -1- TISSUE SPECIFICITY: Testis and brain specific.
CC -1- SIMILARITY: Belongs to the PNMA family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF335505; AAL73196.1; -.
DR EMBL; AF335505; AAL73196.1; -.
KW Nuclear protein.
FT DOMAIN 335 341 POLY-GLU.
FT SEQUENCE 353 AA; 39793 MW; FFC3A717FDCEADA3 CRC64;
SQ
Query Match 33.9%; Score 822; DB 1; Length 353;
Best Local Similarity 49.2%; Pred. No. 2.6e-52;
Matches 174; Conservative 55; Mismatches 117; Indels 8; Gaps 4;
Qy 1 MPTLLQDWCGRGHEHLNTRCMLILGIPEDCGEDFEETLQEAACHLGRYRIGRMFRREE 60
Db 1 MAMTLLDWCGRGMDVNSQALLVWGIPVNCDETIETLQAMPQVS-YRVLGRMFREE 59
Qy 61 NQAAILLELAQDIDYALLPREIPGKGGPWEIVKPRNSDGEFLNRLNRFLEERRTVSDM 120
Db 60 NAKAALLELTGVDYSLIPREMPGKGLKVKVPPTSDAFLERHLFLAREGWTQDV 119
Qy 121 NRVLGSDNTCSAPRTVISPEFTWQAOTLGAAVQLPQLQMLYRELRFVSGNTISIPGALAF 180
Db 120 ARVLGFQNPAPAPGPETPAEMLN--TLDNVIQPLVESIWYKLTFLSGKDIPGGEETF 177
Qy 181 DAWLEHTTEMLQWQVEGEGKRRLEMCRLCPALQVVSGLRASNASITVEECLAQOVF 240
Db 178 DSWLEHSENEIEEWQVSDIEKRRRLMSLGRPADVIRILKTNNAITTAELKALEQVF 237
Qy 241 GPVESHKIAQVKLCAYQEAQEKVSSFFVLRLEPLLRQAVENNVVSRNNVQTRLKRVLSG 300
Db 238 GSVESSRDAQVRFLNTYQNPGEKLSYVIRLEPLQLQKVDKGVIDKONVQARLEQVIAG 297
Qy 301 ATLPDKLRDULKMKQRKRP-PGFLALVKLLRE-----EEWEATIGPDRESLEG 349
Db 298 ANHSGALRRQLWLAGAEGPAPNLFQLLVQIREBEAKKEEAEAAALQLGLEG 351
RESULT 5
PMAL HUMAN STANDARD; PRT; 353 AA.
ID PMAL HUMAN
AC Q8ND90; Q95144; Q8NG07;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Paraneoplastic antigen Mal (Neuron- and testis-specific protein 1)
DE (37 kDa neuronal protein).
GN PNMA1 OR MAL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RC TISSUE=Cerebellum;

```


RA Stapleton M., Soares M.B., Ronaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McQuinn P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
RA Fahney J., Helton E., Kettelman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Mediates caspase-dependent apoptosis.
CC -!- SUBUNIT: Homodimer. Binds BAX, BCL2 and BCLX.
CC -!- TISSUE SPECIFICITY: Ubiquitously expressed with higher levels in
CC heart and brain.
CC -!- DOMAIN: The BH3-like domain is required for association with BAX
CC and for mediating apoptosis. The three BH domains (BH1, BH2, and
CC BH3) of BAX are all required for mediating protein-protein
CC interaction.
CC -!- SIMILARITY: Belongs to the PNMA family.
CC -!- CAUTION: Ref.2 sequence differs from that shown due to a
CC frameshift in position 102.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF305550; AAG31786.1; -
CC EMBL; AK024029; BAB14788.1; ALT_SEQ.
CC EMBL; BC015044; AAH15044.1; -
CC Genbank; HGNC:16658; MOAP1.
CC Apoptosis.
CC SITE 120 127 BH3-LIKE.
CC MUTAGEN 120 127 RESULTING IN A NONAPOPTOTIC PROTEIN.
CC MUTAGEN 120 127 L->E: WEAKENED INTERACTION WITH BAX,
CC MUTAGEN 125 127 GHE->VLA: ABROGATED INTERACTION WITH BAX,
CC MUTAGEN 125 127 RESULTING IN A NONAPOPTOTIC PROTEIN.
CC CONFLICT 244 244 T -> A (IN REF. 2).
CC CONFLICT 258 258 Y -> H (IN REF. 2).
CC CONFLICT 259 259 Q -> H (IN REF. 1).
CC SEQUENCE 351 AA; 39512 MW; 5310142AC02B563C CRC64;
CC -----
CC Query Match 31.7%; Score 768.5; DB 1; Length 351;
CC Best Local Similarity 46.6%; Pred. No. 1.9e-48;
CC Matches 165; Conservative 60; Mismatches 118; Indels 11; Gaps 6;
CC -----
CC 1 MPTLLQDWCRCGEHLNTRRCMLILGIPDCGDEFEFTLQACRHLGRLYRIVGIMFRREE 60
CC 1 MTLRLLEDWCRCGMDNPRKALLIAGISQSCVAEIEEALQAGLAPLGEYRLGLRMFRDE 60
CC 61 NQAAILLELAQDIDYALLPREIPCKGGPWEIVKPRNSDGEFLNRLNRFLEERRTVSDM 120
CC 61 NRKVALVGLTATETSHALVPKGIWVIFKPPDPDNTFLSRNLNFEFLAGEGTMVGL 120
CC 121 NRVLGSDTNCAPRTVISPEFWT--WAQTLGAADVQPLEQMLYRELRFVSGNTISIPGAL 178
CC 121 SRALCHENGSLDPEQGMIPENWAPMLAAL--EALQALQCLYKYLKLVFSRESPEPEEE 179
CC 179 AFDALWEHTTLMQWQVPEGEKRRRLMECLRGALQVQVSLGRASNASITVEECALAQ 238
CC 180 EFGRMWFTTQMIKAWQVPDVEKRRLLSLRGPALDVRVLKINNPLITVDECIQALEE 239
CC 239 VFGPVESHKIAQVCKLCKAYQAGEKVSFVLRLPLQLQRAVNNVSRNNVQTRKLKVL 298

DB 240 VFGVTNPRELQVKLYTYQDEKLSAYVLRLEPLQKLVQRCALIEDAVNQARLDQVI 299
QY 299 SGATLDFKLRDLKLMKQRRKPPGFIALVKLRE---EEWEATLGPDRSLEG 349
DB 300 AGA-VHKTIRRELNL-PEDGAPCFGLQLLVLIKDYAAEEBELL---QALLEG 348
RESULT 7
MOAP1 MACFA STANDARD; PRT; 351 AA.
AC Q95K14;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DE 15-MAR-2004 (Rel. 43, Last annotation update)
DE Modulator of apoptosis 1 (MAP-1).
GN MOAP1.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheciniae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Temporal cortex;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
RT libraries";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Mediates caspase-dependent apoptosis (By similarity).
CC -!- SUBUNIT: Homodimer. Binds BAX, BCL2 and BCLX (By similarity).
CC -!- DOMAIN: The BH3-like domain is required for association with BAX
CC and for mediating apoptosis. The three BH domains (BH1, BH2, and
CC BH3) of BAX are all required for mediating protein-protein
CC interaction (By similarity).
CC -!- SIMILARITY: Belongs to the PNMA family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB060854; BAB46873.1; -
CC Apoptosis.
CC SITE 120 127 BH3-LIKE.
CC MUTAGEN 336 339 POLY-GLU.
CC SEQUENCE 351 AA; 39623 MW; C7530E4496A6FFB3 CRC64;
CC -----
CC Query Match 31.4%; Score 761.5; DB 1; Length 351;
CC Best Local Similarity 46.6%; Pred. No. 6.3e-48;
CC Matches 165; Conservative 59; Mismatches 119; Indels 11; Gaps 6;
CC -----
CC 1 MPTLLQDWCRCGEHLNTRRCMLILGIPDCGDEFEFTLQACRHLGRLYRIVGIMFRREE 60
CC 1 MTLRLLEDWCRCGMDNPRKALLIAGISQSCVAEIEEALQAGLAPLGEYRLGLRMFRDE 60
CC 61 NQAAILLELAQDIDYALLPREIPCKGGPWEIVKPRNSDGEFLNRLNRFLEERRTVSDM 120
CC 61 NRKVALVGLTATETSHALVPKGIWVIFKPPDPDNTFLSRNLNFEFLAGEGTMVGL 120
CC 121 NRVLGSDTNCAPRTVISPEFWT--WAQTLGAADVQPLEQMLYRELRFVSGNTISIPGAL 178
CC 121 SRALCHENGSLDPEQGMIPENWAPMLAAL--EALQALQCLYKYLKLVFSRESPEPEEE 179
CC 179 AFDALWEHTTLMQWQVPEGEKRRRLMECLRGALQVQVSLGRASNASITVEECALAQ 238
CC 180 EFGRMWFTTQMIKAWQVPDVEKRRLLSLRGPALDVRVLKINNPLITVDECIQALEE 239
CC 239 VFGPVESHKIAQVCKLCKAYQAGEKVSFVLRLPLQLQRAVNNVSRNNVQTRKLKVL 298

Db 240 VEGVTNPRELVQKYLTTQKDEKLSAYVLRLEPLLQKLVQGAERDAVNAQRLDQVI 299
 QY 299 SGATLPDKLRLDKLQKQKRPFGFLALVKLLRE---EEWEATLQPDRESLEG 349
 Db 300 AGA-VHKTIRRELNL-PEDGPAPGFLQLLVLIKDYEAABEEALL---QEVLEG 348

RESULT 8

PM2_MOUSE
 ID PM2_MOUSE STANDARD; PRT; 365 AA.
 AC Q8BHK0;
 DT 15-MAR-2004 (Rel. 43, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Paraneoplastic antigen M2 homolog.
 GN PM2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=C57BL/6J; TISSUE=Brain cortex;
 RX MEDLINE=22354683; PubMed=12466851;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojohori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik M., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawai H., Kawasawa Y., Redziarski R.M., King B.L.,
 RA Konggaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagahima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar (By similarity).
 CC -!- SIMILARITY: Belongs to the PMA family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AK043718; BAC31626.1;
 DR EMBL; AK043910; BAC31700.1;
 DR MGD; MGI:2444129; Pma2.
 KW Nuclear protein.
 FT DOMAIN 338 POLY-GLU.
 SEQUENCE 365 AA; 41201 MW; 0F90C940B9D843D9 CRC64;

Query Match

29.9%; Score 725.5; DB 1; Length 365;

Best Local Similarity 43.4%; Pred. No. 2.7e-45;
 Matches 148; Conservative 76; Mismatches 110; Indels 7; Gaps 3;
 QY 1 MPUTLQDWCRCGHELNTRCMLILGIPDCGDEFEETLQEA CRHIGRVRVIGMRFRREE 60
 Db 1 MAVALLEEWCKMGVDVQKSLVVDIPVDCGEBEIQTVLQEA LKCVGSYLLGKIFQKQD 60
 QY 61 NQAAILLELAQDIDYALLPREIPKCGPWEIFKPRNSDGEFLNRLNRLLEERRTVSDM 120
 Db 61 NTSVVLVLMEDTDMSSVPSEVQGGVWKVFKPTNPQDTEFLQRLNLFLEKGGQTVAGM 120
 QY 121 NRVLSGDTNCSAPRVTVISPEFWTQAOTLGAAV----QPLLEOMLYRELRFVSGNTTISIPG 176
 Db 121 FRALKHEGVSPATPPCTSPSEL--LAHLTQAMVHGQRP LL-PVKYCKMIFSGSTAAPE 177
 QY 177 ALAFDAWLHETTEMLOMQVPEGEKRRRLMECLRG PALQVSGLRASNASITVEEC LAAL 236
 Db 178 EEPFEVWLEQATEIAKEWPIPEAEKKRWVAESLRG PALDLMHIVQADNPSISVGECL EAF 237
 QY 237 QOVFGPVE SHKTAOVKLCAYQEA GKVSSFVLRLEPILQRAVENNVVRRNVNQT LKR 296
 Db 238 KOVFGSTESRRTSQVKYLRVYQOGEKISAYVLRLETLRRRAVEKRAIPNTIADQVRLEQ 297
 QY 297 VLSGATLPDKLRDKLQKQKRPFGFLALVKLLREEEWE 337
 Db 298 VMAGANLGNVLCRLQELQKDGFLPTFLQ LMKVIREEEEEE 338

RESULT 9

MOPI_MOUSE
 ID MOPI_MOUSE STANDARD; PRT; 352 AA.
 AC Q9EHH6;
 DT 15-MAR-2004 (Rel. 43, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Modulator of apoptosis 1 (MAP-1).
 GN MOAP1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21264738; PubMed=11060313;
 RA Tan K.O., Tan K.M.L., Chan S.-L., Yee K.S.Y., Bevort M., Ang K.C.,
 RA Yu V.C.;
 RT "MAP-1, a novel proapoptotic protein containing a BH3-like motif that
 RT associates with Bax through its Bcl-2 homology domains.";
 RL J. Biol. Chem. 276:2802-2807(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Frieschmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";

```

RL Nature 409:685-690 (2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6Ncr; TISSUE=Hematopoietic, and Kidney;
RX MEDLINE=22388257; PubMed=12477932;
RA Srausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toehiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abranson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettaman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- FUNCTION: Mediates caspase-dependent apoptosis.
CC -!- SUBUNIT: Homodimer. Binds BAX, BCL2 and BCLX (By similarity).
CC -!- DOMAIN: The BH3-like domain is required for association with BAX
CC and for mediating apoptosis. The three BH domains (BH1, BH2, and
CC BH3) of BAX are all required for mediating protein-protein
CC interaction (By similarity).
CC -!- SIMILARITY: Belongs to the PMA family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
CC EMBL; AF305551; AAC311787.1; -
CC DR EMBL; AK019599; BAB31810.1; -
CC DR EMBL; BC014715; AAK14715.1; -
CC DR EMBL; BC055374; AAH55374.1; -
CC DR MGD; MGI:1915555; Woapl.
CC KW Apoptosis.
FT SITE 120 127 BH3-LIKE.
FT DOMAIN 335 340 POLY-GLU.
FT CONFLICT 57 57 R -> K (IN REF. 3; AAH55374).
FT SEQUENCE 352 AA; 39404 MW; 8F4630D080495D98 CRC64;
Query Match 29.3%; Score 711; DB 1; Length 352;
Best Local Similarity 42.2%; Pred. No. 2.8e-44;
Matches 151; Conservative 70; Mismatches 119; Indels 18; Gaps 7;
QY 1 MPTLLQWCNGEHLNTRCMLIGIPDCGDEBFEETLOACRHLYRGVIGRMFRREE 60
Db 1 MTLLEDCWCGRMDNPRKALVAGIPTTCGVADIEEALQAGLAPLGEHLLGRMFRDE 60
QY 61 NQAAILLELAQDIDYALLPREIPCKGGPWEIVKPRNSDGFELNRLNFFLEERRTVSDM 120
Db 61 NKNVALIGLTVETGSALVPKEIPAKGGVMRVIFKPPDTSDFLCRLNFEFLKGGTGMGEL 120
QY 121 NRVLGS-----DTNCSAPRTVISPEFTWTAQTLGAAVQPLLEOMLYRELRFVSGNTIS 173
Db 121 TRVLGNRNDPLGLDPGIMPEIR-APML--AAQALNEALKTLQVLYKLVSVSGRDP 176
QY 174 IPGALAFDAWLEHTTEMLOMQVPEGEKRRKLMECLRGPAQVVSGLRASNASITVECL 233
Db 177 GPGEEFESWFMHTSQVMKMTQVSDVEXKRRLLIESLRGPAFEIIRVLKINPNFFITVABCL 236
QY 234 AALQVQFQVESHKTAQVKLCKAQEAGEKVSFVLRLEPLLQRAVENNVVSRNVNQR 293

```

```

Db 237 KTLTETFGIINDPRAQVKLTYYQTKDEKLSAVLELEFLLQKLVKGALKEVNNQAR 296
QY 294 LKEVLGATLPDKRLDKLKMQRKPPGFLALVKLL--REEEWEATLPGDRESLEG 349
Db 297 LDQVIAGA-VHKSVRRELGL-PEGSPAPGLQLTLTKKEAEEVLL---QAELEG 349

RESULT 10
ID LAM3 MOUSE STANDARD; PRT; 474 AA.
AC P48680;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Lamin B3.
GN LMNB2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93154351; PubMed=8094052;
RA Furukawa K., Hotta Y.;
RT "cDNA cloning of a germ cell specific lamin B3 from mouse
RT spermatocytes and analysis of its function by ectopic expression in
RT somatic cells.";
RL EMBO J. 12:97-106 (1993).
CC -!- FUNCTION: Lamins are components of the nuclear lamina, a fibrous
CC layer on the nucleoplasmic side of the inner nuclear membrane,
CC which is thought to provide a framework for the nuclear envelope
CC and may also interact with chromatin.
CC -!- SUBCELLULAR LOCATION: Nucleoplasmic side of the inner nuclear
CC membrane.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=B3;
CC IsoId=P48680-1; Sequence=Displayed;
CC Name=B2;
CC IsoId=P21619-1; Sequence=External;
CC Note=No experimental confirmation available;
CC -!- TISSUE SPECIFICITY: Germ cell specific.
CC -!- PTM: B-TYPE LAMINS UNDERGO A SERIES OF MODIFICATIONS, SUCH AS
CC FARNESYLATION AND PHOSPHORYLATION. INCREASED PHOSPHORYLATION OF
CC THE LAMINS OCCURS BEFORE ENVELOPE DISINTEGRATION AND PROBABLY
CC PLAYS A ROLE IN REGULATING LAMIN ASSOCIATIONS.
CC -!- MISCELLANEOUS: The structural integrity of the lamina is strictly
CC controlled by the cell cycle, as seen by the disintegration and
CC formation of the nuclear envelope in prophase and telophase,
CC respectively.
CC -!- SIMILARITY: Belongs to the intermediate filament family. THIS IS A
CC B TYPE LAMIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
CC EMBL; D13455; BAA02708.1; -.
CC PIR; S28419; S28419.
CC MGD; MGI:96796; Lmbb2.
CC GO; GO:0005638; C:lamin filament; IDA.
CC InterPro; IPR001664; IF.
CC InterPro; IPR001322; IF_tail_C.
CC Pfam; PF00038; Filament; 1.
CC Pfam; PF00932; IF_tail; 1.
CC PROSITE; PS00226; IF; 1.
CC Intermediate filament; Coiled coil; Nuclear protein; Lipoprotein;
CC Prenylation; Phosphorylation; Alternative splicing.
KW

```


OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 [1]
 RN SEQUENCE FROM N.A.
 RC TISSUE=Cerebellum, Retina, and Spinal cord;
 RX MEDLINE=20347255; PubMed=10764728;
 RA Stabach P.R.; Morrow J.S.;
 RT "Identification and characterization of beta V spectrin, a mammalian
 ortholog of Drosophila beta H spectrin.";
 RL J. Biol. Chem. 275:21385-21395(2000).
 CC -!- SUBUNIT: Probably associates with an alpha chain.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic. Detected prominently in the
 outer segments of photoreceptor rods and cones and in the
 basolateral membrane and cytosol of gastric epithelial cells.
 CC -!- TISSUE SPECIFICITY: Expressed at very low levels in many tissues,
 with strongest expression in cerebellum, spinal cord, stomach,
 pituitary gland, liver, pancreas, salivary gland, kidney, bladder,
 and heart.
 CC -!- SIMILARITY: Belongs to the spectrin family.
 CC -!- SIMILARITY: Contains 2 calponin-homology (CH) domains.
 CC -!- SIMILARITY: Contains 1 PH domain.
 CC -!- SIMILARITY: Contains 31 spectrin repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 DR EMBL; AF233523; AAF65317.1; -.
 DR HSPSP; Q01082; 1BKX.
 DR Genew; HGNC:15680; SPTBN5.
 DR MIM; 605916; -.
 DR GO; GO:0016020; C:membrane; NAS.
 DR GO; GO:0008091; C:spectrin; NAS.
 DR GO; GO:0003779; F:actin binding; NAS.
 DR GO; GO:0030036; P:actin cytoskeleton organization and biogenesis; NAS.
 DR InterPro; IPR001589; Actbind actinin.
 DR InterPro; IPR001715; Calponin-like.
 DR InterPro; IPR001849; PH.
 DR InterPro; IPR002017; Spectrin.
 DR InterPro; IPR001605; Spectrin_PH.
 DR Pfam; PF00307; CH; 2.
 DR Pfam; PF00169; PH; 1.
 DR Pfam; PF00435; spectrin; 30.
 DR PRINTS; PR00683; SPECTRINPH.
 DR SMART; SM00033; CH; 2.
 DR SMART; SM00233; PH; 1.
 DR SMART; SM00150; SPEC; 29.
 DR PROSITE; PS00019; ACTININ_1; 1.
 DR PROSITE; PS00020; ACTININ_2; FALSE_NEG.
 DR PROSITE; PS00021; CH; 2.
 DR PROSITE; PS50003; PH DOMAIN; 1.
 KW Cytoskeleton; Membrane; Repeat; Actin-binding; Actin capping.
 FT DOMAIN 1 279 ACTIN-BINDING (BY SIMILARITY).
 FT DOMAIN 54 159 CH 1.
 FT DOMAIN 177 279 CH 2.
 FT REPEAT 306 416 SPECTRIN 1.
 FT REPEAT 426 531 SPECTRIN 2.
 FT REPEAT 533 637 SPECTRIN 3.
 FT REPEAT 639 743 SPECTRIN 4.
 FT REPEAT 745 810 SPECTRIN 5.
 FT REPEAT 867 894 SPECTRIN 6.
 FT REPEAT 896 998 SPECTRIN 7.
 FT REPEAT 1102 1206 SPECTRIN 8.
 FT REPEAT 1208 1312 SPECTRIN 9.
 FT REPEAT 1314 1417 SPECTRIN 10.
 FT REPEAT 1419 1488 SPECTRIN 11.
 FT REPEAT 1520 1624 SPECTRIN 12.
 FT REPEAT 1626 1728 SPECTRIN 13.
 FT REPEAT 1730 1835 SPECTRIN 14.

FT REPEAT 1837 1941 SPECTRIN 15.
 FT REPEAT 1943 2047 SPECTRIN 16.
 FT REPEAT 2049 2147 SPECTRIN 17.
 FT REPEAT 2149 2253 SPECTRIN 18.
 FT REPEAT 2255 2301 SPECTRIN 19.
 FT REPEAT 2314 2362 SPECTRIN 20.
 FT REPEAT 2364 2468 SPECTRIN 21.
 FT REPEAT 2470 2574 SPECTRIN 22.
 FT REPEAT 2576 2680 SPECTRIN 23.
 FT REPEAT 2682 2785 SPECTRIN 24.
 FT REPEAT 2787 2891 SPECTRIN 25.
 FT REPEAT 2893 2997 SPECTRIN 26.
 FT REPEAT 2999 3103 SPECTRIN 27.
 FT REPEAT 3105 3210 SPECTRIN 28.
 FT REPEAT 3212 3312 SPECTRIN 29.
 FT REPEAT 3314 3418 SPECTRIN 30.
 FT REPEAT 3420 3482 SPECTRIN 31.
 FT DOMAIN 3533 3641 PH.
 SQ SEQUENCE 3674 AA; 416832 MW; 8C0FDFAD0DD0C7C18 CRC64;

Query Match 4.7%; Score 113; DB 1; Length 3674;
 Best Local Similarity 21.0%; Pred. No. 12;
 Matches 93; Conservative 63; Mismatches 155; Indels 132; Gaps 22;

Qy 36 EETLQEAQRHLGRVYVIGRMFRRENAQAAIL-----ELAQDIDYALLPR----- 80
 Db 1332 EKLMAAHEFGARRNLIQTLKRHEAAESELLATRRHVEALQVGRLLSRPCGQEDIQ 1391
 Qy 81 -EIPGKGPMVEVIVKPRNSDGEFL-----NRLNRFLEERVTYSDMNRVL-----GSDT 128
 Db 1392 TRLQGLRSKWEALNKNKWTGDELQQAQGBQLLRLQDAKEQLEQLEGALQSSETGDL 1451
 Qy 129 NCSAPRTVISPEFTWQTL-----GAAVQPLLE--OMLYRELRFVSGNTIS 173
 Db 1452 RSSQRLQKRHQQLSESRITLAAKMAALASMAHGAASPAILEETQKHLRLELQGH-LA 1510
 Qy 174 IPGALAFDAWLE-----HTTEMLQW--OVPEGEKRRRLMECLRG-----PALQV- 216
 Db 1511 IRG-LQQAQVELHQFCHLSNMELSWAEHMPHSP-TSYTECLNGAOSLHRKHKLQVE 1568
 Qy 217 -----VSGLRASNASIT-----VBECLAAALQVFGPVESHKIAQVKLCKAYOE 259
 Db 1569 VKAHQGVQVRLSSGRSLAASGHFQAQHVQEC-----QELGHWAEALERACEARAQ 1620
 Qy 260 AGEKVSSF-----VLRLE-----PLL-----ORAVENNVSRNVNQTRL----- 294
 Db 1621 CLOQAVTFQYFLDVSELEGWVEEKRPLVSSRDYGRDEAATRLINKHQAQOEELAIYWS 1680
 Qy 295 -----KRVLSGATLP-----DKLRDKLKMKO-----RRKPPGFLALVKLLREEE 334
 Db 1681 SMEELDQTAQTLTGPEVPEQQRVVQERLREQRLALQELAAATRELEGTLLRHEFLREAE 1740
 Qy 335 EWEATLGPDRSLEGLVAPRPP 357
 Db 1741 DLQGWLASQKQAKGESLGEDP 1763

RESULT 13
 CK13 HUMAN
 ID CK13 HUMAN STANDARD; PRT; 373 AA.
 AC Q02833;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Protein C11orf13 (HRAS1-related cluster protein 1).
 GN C11orf13 OR HRC1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE=Placenta;

```
RX MEDLINE=93052330; PubMed=1339391;
RA Weitzel J.N., Kasperczyk A., Mohan C., Krontiris T.G.;
RT "The HRA51 gene cluster: two upstream regions recognizing transcripts
RL and a third encoding a gene with a leucine zipper domain.";
CC Genomics 14:309-319(1992).
CC -I- ALTERNATIVE PRODUCTS.
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q02833-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q02833-2; Sequence=VSP_004136, VSP_004137;
CC -I- SIMILARITY: Contains 1 Ras-associating domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M91083; AA58667.1; --
CC PIR; A44478; A44478.
CC PIR; B44478; B44478.
CC Genew; HGNC:1166; C11orf13.
CC MIM; 143023; --
CC GO; GO:0005634; C:nucleus; NAS.
CC GO; GO:0003677; F:DNA binding; NAS.
CC GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
CC InterPro; IPR000159; RA_domain.
CC Pfam; PF00788; RA; 1.
CC SMART; SM00314; RA; 1.
CC PROSITE; PS50200; RA; 1.
CC Coiled coil; Alternative splicing.
CC DOMAIN 6 89 RAS-ASSOCIATING.
CC PRO-RICH.
CC FT DOMAIN 130 149 COILED COIL (POTENTIAL).
CC FT DOMAIN 175 227 COILED COIL (POTENTIAL).
CC FT DOMAIN 228 236 PRO-RICH.
CC FT DOMAIN 248 297 COILED COIL (POTENTIAL).
CC FT DOMAIN 304 322 PRO-RICH.
CC FT VARSPLIC 318 337 GLPAPPRESILGAPESHA -> VGVVLGGGWGVTWPSP
CC TFS (in isoform 2).
CC /FTID=VSP_004136.
CC Missing (in isoform 2).
CC /FTID=VSP_004137.
CC SEQUENCE 373 AA; 39945 MW; 55DF1DD81C37CAF3 CRC64;
Query Match
Best Local Similarity 4.5%; Score 110; DB 1; Length 373;
Matches 97; Conservative 49; Mismatches 144; Indels 110; Gaps 20;
QY 22 LTLGIPEDCGEDEFETLQEAACRHLYGRVYIGRMFRER-NAQAALLELAQD 72
DQ 20 VCVGVSEQTCTCEVVIALQAIGTGRVIVLORLEKREQLLPQECPCVQAATCGQFASD 79
QY 73 IDYALLPREIPKGG-----PWE-----VIVKPNSDGEFLNRLNREFLEBERRTV 117
DQ 80 VQF-VLRRTGSLAGRPSSDCPPPERCLIRASLPVKPRAALG-----CEPRKTL 128
QY 118 SDMNRVLGSDINCSAPRTISPEFTWATLGAAYQPLLEQMLYRELRFVSGNTISIPGA 177
DQ 129 TPEAPSLRSGPAAP-VPTPGCTDLRGLRLVQRNAEEL-----GH 171
QY 178 LAFDAWLEHTTLMQWQVPEGEKRRRLMECLRGPAQVLSGLASNA---SITVEECIA 234
DQ 172 EAF--W-----SQELRREQAREEGARL-QALSAETAHAARLQALDAQARALELQLA 224
QY 235 ALQOVFGP-----VESHKIAQVK-----LCKAYQAGEKVSFVLRLPEL 274
DQ 225 A--EAPGPPSPMASATERLHODLAVQERQSAEVQSGSLAVSRALEAERA-----LQAQ 276
QY 275 LQRAVENNVSRNNVOTRLKVL--SGATLPDKLRDKLKMQRKPGPLALVKLIRE 332
```

```
Db 277 AQELEELN---RELRCNLQFIQTGAALPPRPRED-----RGPPCTQGLPPARE 325
QY 333 EEWENATLGPDRSLGLEVAVPR---PPARITGVGAVPLP 369
DQ 326 ----ESLLGAPSESHAGAPRPRGGRPHDAELLEVAAPAP 361
RESULT 14
PERQ_HUMAN STANDARD; PRT; 817 AA.
ID PERQ_HUMAN AC 075420;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE PERQ amino acid rich with GYF domain protein 1.
GN PERQ1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99018118; PubMed=9799793;
RA Gloeckner G., Scherer S., Schattevoy R., Boright A.P., Weber J.,
RA Tsui L.-C., Rosenthal A.;
RT "Large-scale sequencing of two regions in human chromosome 7q22:
RT analysis of 650 kb of genomic sequence around the EPO and CUTL1 loci
RT reveals 17 genes.";
RL Genome Res. 8:1060-1073(1998).
CC -I- SIMILARITY: Contains 1 GYF domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF053356; AAC78792.1; --
CC Genew; HGNC:9126; PERQ1.
CC InterPro; IPR003169; GYF.
CC Pfam; PF02213; GYF; 1.
CC SMART; SM00444; GYF; 1.
CC PROSITE; PS50829; GYF; 1.
CC DOMAIN 352 400 GYF.
CC FT DOMAIN 426 495 GLN-RICH.
CC FT DOMAIN 113 120 POLY-GLU.
CC FT DOMAIN 176 181 POLY-GLY.
CC FT DOMAIN 256 260 POLY-GLU.
CC FT DOMAIN 474 480 POLY-PRO.
CC FT DOMAIN 481 486 POLY-GLN.
CC FT DOMAIN 751 757 POLY-GLN.
CC SEQUENCE 817 AA; 89740 MW; C727AD2E7C2E2581 CRC64;
Query Match
Best Local Similarity 4.4%; Score 107.5; DB 1; Length 817;
Matches 101; Conservative 60; Mismatches 119; Indels 195; Gaps 26;
QY 25 GIPEDCGEDEFETLQEAACRHLYGRVYIGRMFRERENAQAALLELAQIDYALLPREI-- 82
DQ 200 GLPEWCLDDEDEE-----MGTF-----DASGAFLEPLKLF 229
QY 83 -----PGKGGWEVIVKPNSDGEFL---NRLNREFLEBERRTVSDMNRVLGSDINCSAPR 134
DQ 230 APHPGFMQKGRKPIPEEQELDFQGLEEEBEPFSEGLEBEGPEADDIR-----G 277
QY 135 VTISPEFTWATLGAAYQPLLEQMLYRELRFVSGNTISIPGALAFDAWLEHTTLMQW 194
DQ 278 IQLSP-----GVSSAGP-----FGDLEDGLKHLQOGIQ-- 308
QY 195 QVPEGEKRRRLMECLRGPAQ-----VVSGLRASNASITVECLAAQVVF----- 240
```



```

Db 309 ---EAEK---LVASLDSSLEEBEFTAAQTQGLRHSAATAPLPLSHGAARKWFYKDPQG 362
QY 241 ---GPVESHKIAQ---VKL-C-KAYQEAQGVSSVFLRL----- 271
Db 363 EIQQPFTTQEAWEAFQAGYFMSLLVKGCDGEGQPLGE---VKNWGRVPPAGPSP 417
QY 272 EPLLQRAVENNVSRNNQTRLKR--VLSGATLPDKLRK--LKIMKQRKPPGFLALV 327
Db 418 PPLL-----GNMDOERLKKQQLAAALYQQLQHQQFLQVSSRQLPQ----- 460
QY 328 KILREEEWEATLGPDRLESLGLEVAVRPP-----ARITGVGAVPLPASGNSFD 376
Db 461 CALRE-----KAALG-----DLTPPPPPPPQQQQQLTAFLOQLALXPPSS----- 502
QY 377 ARPSQGYRRRRGRGQRRGGVARAGSRGSRKRKHTFCYSGEDGHIRVQCINPS 431
Db 503 -RSAEKSSGRS-GRKNASVYRSASSSRRSR-----SGRRRKSCFGAS 547

RESULT 15
SSB DEIRA
ID SSB DEIRA STANDARD; PRT; 301 AA.
AC Q9RY51:
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Single-strand binding protein (SSB) (Helix-destabilizing protein).
GN SSB OR DR0099.
OS Deinococcus radiodurans.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=R1 / ATCC 13939 / DSM 20539 / NCIB 9279;
RA Witte G., Urbanke C., Curth U.;
RT "Single-stranded DNA binding protein from Deinococcus radiodurans -
Cloning, expression and biophysical characterization.";
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=R1 / ATCC 13939 / DSM 20539 / NCIB 9279;
RA Egginton J.M., Wood E.A., Cox M.M.;
RT "The Deinococcus radiodurans ssb gene encodes a contiguous ORF for
single-stranded DNA-binding protein.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
RN (3)
RP SEQUENCE FROM N.A.
RC STRAIN=R1 / ATCC 13939 / DSM 20539 / NCIB 9279;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
radiodurans R1.";
RL Science 286:1571-1577(1999).
CC -!- FUNCTION: This protein is essential for replication of the
chromosome. It is also involved in DNA recombination and repair
(BY similarity).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SIMILARITY: Contains 2 SSB domains.
CC -!- CAUTION: Ref.3 sequence differs from that shown due to two
frameshifts in positions 93 and 169.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial

```

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ564860; CAD92322.1; -.
DR EMBL; AY293617; AAQ18705.1; -.
DR EMBL; AE001873; AAF09692.1; ALT_FRAME.
DR EIR; A75559; A75559.
DR HSP; P02339; IKAW.
DR TIGR; DR0099; -.
DR InterPro; IPR008994; Nucleic acid OB.
DR InterPro; IPR000424; SSB_protein.
DR Pfam; PF00436; SSB; 1.
DR PROSITE; PS00935; SSB; 2.
KW DNA-binding; DNA repair; DNA replication; Complete proteome; Repeat.
FT DOMAIN 5 110
FT DOMAIN 129 232
FT SEQUENCE 301 AA; 32722 MW; 21B2090C970E1B0E CRC64;
SQ
Query Match 4.4%; Score 106.5; DB 1; Length 301;
Best Local Similarity 25.5%; Pred. No. 1.3;
Matches 80; Conservative 39; Mismatches 102; Indels 93; Gaps 21;
QY 116 TVSDMRVLGSD-TNCSAP---RVTI--SPEFTWQTLGAAVQPLLEQMLYRELRFVSG 169
Db 33 TVAGEDRVIGNDGRERNLPWVHRVSVILGKPAEW-----QAERNLKGK 74
QY 170 NTSIPGALLAFDAWLEHTTEMLQWQVPEGEKRRRL-MECLR-----GPAQVW----SGL 220
Db 75 DAVVVEGTLEY-----RQWEAPEGGRKSAVNVKALRMBQLGTPELIQDAGGV 123
QY 221 RASNASITVEECLAALQOVFGPVBESHKIAQVKLCKAYQAGEKYSFVRLLEPLLQRAVE 280
Db 124 RMSGA---MEVL-----VLGNVT--RDPEIR---YTPGDAVL-----LSIAVN 161
QY 281 NNVSRRNVNQLRLKRVLSGATLPDKLRDKLKLKQRRKPPGFALVKLREBEWEATL 340
Db 162 ENYQDRQGQKEKVHYI--DATL---NRDLAENMKELRKGPVIMGLV--NEGWTDKD 214
QY 341 GPDRESLEGLVAPRPPPARITGVG-----AVPLPASGN-----SFDARP-SQGYRRRRGR 389
Db 215 GNKENSTR--VEATRVLEALFAGNAGNANGYAAATPAAPRTQTATASSAARPTSGGYQSQPSR 272
QY 390 GQHRGGVARAGSR 403
Db 273 -----AANTGSR 279

```

Search completed: September 27, 2004, 17:08:48
Job time : 18.1307 secs

Blank

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 27, 2004, 17:05:06 ; Search time 58.2533 Seconds
(without alignments)
2507.755 Million cell updates/sec

Title: US-10-037-860-13

Perfect score: 2423
Sequence: 1 MPTLLQDWCGRGHEHLNTRRC.....VESGNGNWAWDKSHPKSKAK 463

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPREMBL 25.*

1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2423	100.0	463	Q9UL41	Q9ul41 homo sapien
2	2283.5	94.2	455	Q9H0A4	Q9h0a4 homo sapien
3	1796.5	74.1	466	Q8JZW8	Q8jzw8 mus musculu
4	883.5	36.5	448	Q8NET3	Q8net3 homo sapien
5	883.5	36.5	452	Q96PV4	Q96pv4 homo sapien
6	836.5	34.5	364	Q94959	Q94959 homo sapien
7	826.5	34.1	364	Q9GMU3	Q9gm3 macaca fasc
8	824	34.0	353	Q9CYP2	Q9cyp2 mus musculu
9	822	33.9	353	Q8VH24	Q8vh24 rattus norv
10	822	33.9	353	Q8C1C8	Q8c1c8 mus musculu
11	818	33.8	353	Q8NG07	Q8ng07 homo sapien
12	816	33.7	353	Q95144	Q95144 homo sapien
13	793	32.7	150	Q9BE35	Q9be35 macaca fasc
14	768.5	31.7	351	Q96BY2	Q96by2 homo sapien
15	768.5	31.7	399	Q96A40	Q96a40 homo sapien
16	763.5	31.5	351	Q9HAS1	Q9has1 homo sapien

17	761.5	31.4	351	6	Q95KI4	Q95ki4 macaca fasc
18	725.5	29.9	355	11	Q8BHK0	Q8bhk0 mus musculu
19	711	29.3	352	11	Q9ERH6	Q9erb6 mus musculu
20	618.5	25.5	283	4	Q9UL42	Q9ul42 homo sapien
21	470.5	19.4	402	11	Q9CZA5	Q9cza5 mus musculu
22	454.5	18.4	237	4	Q9H833	Q9h833 homo sapien
23	449.5	18.6	194	4	Q95145	Q95145 homo sapien
24	425.5	17.6	192	4	Q8ND90	Q8nd90 homo sapien
25	423.5	17.5	403	4	Q8TE36	Q8te36 homo sapien
26	415.5	17.1	402	4	Q8NIC1	Q8nic1 homo sapien
27	390.5	16.1	393	11	Q9DB17	Q9db17 mus musculu
28	390.5	16.1	393	11	Q8VD24	Q8vd24 mus musculu
29	351.5	14.5	378	4	Q8N3H4	Q8n3h4 homo sapien
30	351.5	14.5	435	4	Q86V59	Q86v59 homo sapien
31	339.5	14.0	430	11	Q80VM8	Q80vm8 mus musculu
32	333.5	13.8	246	11	Q8C533	Q8c533 mus musculu
33	323	13.3	149	4	Q9UL43	Q9ul43 homo sapien
34	202	8.3	327	11	Q8VC32	Q8vc32 mus musculu
35	191	7.9	538	4	Q9H0W5	Q9h0w5 homo sapien
36	191	7.9	538	4	Q8TB26	Q8tb26 homo sapien
37	156	6.4	1394	10	Q8H8L9	Q8h8l9 oryza sativ
38	156	6.4	1433	10	Q942G3	Q942g3 oryza sativ
39	149	6.1	1715	10	Q8H870	Q8h870 oryza sativ
40	148	6.1	1661	10	Q7XBQ6	Q7xbq6 oryza sativ
41	145	6.0	1594	10	Q8LN00	Q8ln00 oryza sativ
42	144.5	6.0	1449	10	Q8SK3	Q8sk3 oryza sativ
43	142.5	5.9	1742	10	Q7XLU2	Q7xlu2 oryza sativ
44	142	5.9	1466	10	Q7XGE5	Q7xge5 oryza sativ
45	142	5.9	1616	10	Q7XW87	Q7xw87 oryza sativ

ALIGNMENTS

RESULT 1

Q9UL41 PRELIMINARY; PRT; 463 AA.
 ID Q9UL41
 AC Q9UL41, 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Paraneoplastic neuronal antigen MA3.
 GN MA3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Gultekin S.H., Voltz R., Rosenfeld M.R., Gerstner E., Eichen J., Posner J.B., Dalmay J.;
 RA "Identification of a novel cancer testis brain antigen using serum antibodies from patients with testicular tumors and paraneoplastic limbic encephalomyelitis";
 RT Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AF083116; AAF05627.1; --
 DR EMBL; HGNC:18742; FNMA3.
 DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR InterPro; IPR001878; Znf CCHC.
 DR Pfam; PF00098; zf-CCHC; 1.
 DR PRINTS; PR00939; C2HCZNFINGER.
 DR SMART; SM00343; Znf_C2HC; 1.
 DR PROSITE; PS50158; ZF_CCHC; 1.
 SQ SEQUENCE 463 AA; 52376 MW; 0843EF6601D8557C CRC64;

Query Match 100.0%; Score 2423; DB 4; Length 463;
 Best Local Similarity 100.0%; Pred. No. 3.2e-185;
 Matches 463; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPTLLQDWCGRGHEHLNTRRCMLILGIPDCGDEFEETLQACRHLGRYVIGRMFRREE 60
 |||||

Db 1 MPTLLQDWCGRGHEHLNTRRCMLILGIPDCGDEFEETLQACRHLGRYVIGRMFRREE 60
 |||||

```

QY 61 NAQAIILELAQDIDYALLPREIFGKGGPWEVIVKPRNSDGEFLNRLNRFLEEBERRTVSDM 120
Db 61 NAQAIILELAQDIDYALLPREIFGKGGPWEVIVKPRNSDGEFLNRLNRFLEEBERRTVSDM 120
QY 121 NRVLGSDTNCSPRVTTISPEFWTWAQTLGAAVQPLLEQMLYRELRFVSGNTTISIPGALAF 180
Db 121 NRVLGSDTNCSPRVTTISPEFWTWAQTLGAAVQPLLEQMLYRELRFVSGNTTISIPGALAF 180
QY 181 DAWLEHTTEMLQWQVPEGEKRRRLMECLRGPALQVVGSLGRASNASTIVTECLAAALQOVF 240
Db 181 DAWLEHTTEMLQWQVPEGEKRRRLMECLRGPALQVVGSLGRASNASTIVTECLAAALQOVF 240
QY 241 GPVESHKIAQVLCCKAQEAGEKVSFVLRLPLELQRAVENNVSRNNVQTRLKRVLG 300
Db 241 GPVESHKIAQVLCCKAQEAGEKVSFVLRLPLELQRAVENNVSRNNVQTRLKRVLG 300
QY 301 ATPDLKRLDKLKMQRKPPGFLALVKLLREEEWEATLGPDRSLEGLVAPRPPARI 360
Db 301 ATPDLKRLDKLKMQRKPPGFLALVKLLREEEWEATLGPDRSLEGLVAPRPPARI 360
QY 361 TGVGAVPLPASGNSFDARPSQGVRRRRGRGQHRGGVARAGSGRSKRKHHTFCYSCGD 420
Db 361 TGVGAVPLPASGNSFDARPSQGVRRRRGRGQHRGGVARAGSGRSKRKHHTFCYSCGD 420
QY 421 GHIRVQCINPSNLLLVKQKQAAVESGNGWMDKSHPKSKAK 463
Db 421 GHIRVQCINPSNLLLVKQKQAAVESGNGWMDKSHPKSKAK 463

```

RESULT 2

```

Q9H0A4 PRELIMINARY; PRT; 455 AA.
ID Q9H0A4
AC Q9H0A4;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN DKFZP344K25.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=21154917; PubMed=11230166;
RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
RA Ansoerge W., Boecker M., Bloeker H., Bauersachs S., Blum H.,
RA Lauber J., Duesterhoeft A., Beyer A., Koehler K., Strack N.,
RA Mewes H.W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,
RA Wambutt R., Korn B., Klein M., Poustka A.;
RA "Towards a Catalog of Human Genes and Proteins: Sequencing and
RT Analysis of 500 Novel Complete Protein Coding Human cDNAs.";
RL Genome Res. 11:422-435 (2001).
DR EMBL; ALI36878; CAB66812.1; -.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR001878; Znf.CCHC.
DR Pfam; PF00098; zf-CCHC; 1.
DR PRINTS; PR00939; C2HCZNFINGER.
DR SMART; SM00343; Znf.C2HC; 1.
DR PROSITE; PS50158; ZF_CCHC; 1.
KW Hypothetical protein.
SQ SEQUENCE 455 AA; 51514 MW; 9AC8CC06685A8956 CRC64;

```

```

Query Match 94.2%; Score 2283.5; DB 4; Length 455;
Best Local Similarity 98.2%; Pred. No. 4.5e-174;
Matches 439; Conservative 3; Mismatches 4; Indels 1; Gaps 1;
QY 1 MPTLLQDWCRCGEHLNTRCMLILGIPDCGDEFEETLQEAACHLGRYRVIGRMFRREE 60
Db 1 MPTLLQDWCRCGEHLNTRCMLILGIPDCGDEFEETLQEAACHLGRYRVIGRMFRREE 60
QY 61 NAQAIILELAQDIDYALLPREIFGKGGPWEVIVKPRNSDGEFLNRLNRFLEEBERRTVSDM 120

```

```

Db 61 NAQAIILELAQDIDYALLPREIFGKGGPWEVIVKPRNSDGEFLNRLNRFLEEBERRTVSDM 120
QY 121 NRVLGSDTNCSPRVTTISPEFWTWAQTLGAAVQPLLEQMLYRELRFVSGNTTISIPGALAF 180
Db 121 NRVLGSDTNCSPRVTTISPEFWTWAQTLGAAVQPLLEQMLYRELRFVSGNTTISIPGALAF 180
QY 181 DAWLEHTTEMLQWQVPEGEKRRRLMECLRGPALQVVGSLGRASNASTIVTECLAAALQOVF 240
Db 181 DAWLEHTTEMLQWQVPEGEKRRRLMECLRGPALQVVGSLGRASNASTIVTECLAAALQOVF 240
QY 241 GPVESHKIAQVLCCKAQEAGEKVSFVLRLPLELQRAVENNVSRNNVQTRLKRVLG 300
Db 241 GPVESHKIAQVLCCKAQEAGEKVSFVLRLPLELQRAVENNVSRNNVQTRLKRVLG 300
QY 301 ATPDLKRLDKLKMQRKPPGFLALVKLLREEEWEATLGPDRSLEGLVAPRPPARI 360
Db 301 ATPDLKRLDKLKMQRKPPGFLALVKLLREEEWEATLGPDRSLEGLVAPRPPARI 360
QY 361 TGVGAVPLPASGNSFDARPSQGVRRRRGRGQHRGGVARAGSGRSKRKHHTFCYSCGD 420
Db 361 TGVGAVPLPASGNSFDARPSQGVRRRRGRGQHRGGVARAGSGRSKRKHHTFCYSCGD 420
QY 421 GHIRVQCINPSNLLLVKQKQAAVESG 447
Db 421 GHIRVQCINPSNLLLVKQKQAAVESG 446

```

RESULT 3

```

Q8JZW8 PRELIMINARY; PRT; 466 AA.
ID Q8JZW8
AC Q8JZW8;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to paraneoplastic antigen MA3.
GN PMA3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RA Strausberg R.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC036726; AAH36726.1; -.
DR MGD; MGI:2180565; Pma3.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001544; Aminotrans IV.
DR InterPro; IPR001878; Znf.CCHC; 1.
DR Pfam; PF00098; zf-CCHC; 1.
DR PRINTS; PR00939; C2HCZNFINGER.
DR PROSITE; PS50158; ZF_CCHC; 1.
SQ SEQUENCE 466 AA; 54040 MW; 113787E37B0B3AAD CRC64;

```

```

Query Match 74.1%; Score 1796.5; DB 11; Length 466;
Best Local Similarity 73.4%; Pred. No. 4e-135;
Matches 345; Conservative 49; Mismatches 65; Indels 11; Gaps 4;
QY 1 MPTLLQDWCRCGEHLNTRCMLILGIPDCGDEFEETLQEAACHLGRYRVIGRMFRREE 60
Db 1 MPTLLQDWCRCGEHLNTRCMLILGIPDCGDEFEETLQEAACHLGRYRVIGRMFRREE 60
QY 61 NAQAIILELAQDIDYALLPREIFGKGGPWEVIVKPRNSDGEFLNRLNRFLEEBERRTVSDM 120
Db 61 NAQAIILELAQDIDYALLPREIFGKGGPWEVIVKPRNSDGEFLNRLNRFLEEBERRTVSDM 120
QY 121 NRVLGSDTNCSPRVTTISPEFWTWAQTLGAAVQPLLEQMLYRELRFVSGNTTISIPGALAF 180
Db 121 NRVLGSDTNCSPRVTTISPEFWTWAQTLGAAVQPLLEQMLYRELRFVSGNTTISIPGALAF 180

```

```
181 DAWLEHTTEMLQMWQVPEGEKRRRLMECLRGPAQVVGSLRAGNASITVECLAAQOVF 240
181 DSWLEHTTEMLQMWQVPEGEKRRRLMECLRGPAQVVGSLRAGNASITVECLAAQOVF 240
241 GPVESHKIAQVKLCAYQAEAGKVSFVLRLPQLQRAVENNVSRNNVOTRLKRVLSG 300
241 GSDVNRKIAQLKFCAYQAEAGKVSFVLRLPQLQRAVENNVSRNNVOTRLKRVLSG 300
301 ATLPLDKLRDLKLMKQRRKPPGFLALVKLLR-BEEWEATLGPDRSLEGLVAPRPPAR 359
301 AILSAKUREKLMKQRRKPPGFLALVKLLR-BEEWEATLGPDRSLEGLVAPRPPAR 359
360 ITGVGA-----VPLPASGNSFDARPSQGYRR--RGRGHRGGVARGSGRSRKRKHTF 413
360 ITGVGA-----VPLPASGNSFDARPSQGYRR--RGRGHRGGVARGSGRSRKRKHTF 413
358 -SNIGSEERELFVPAFGSVLEERPYQSGRRRRHRRGQHRKGGVPRDSSQGRKQNYDTF 416
414 CYSCEGDGHRVQCINPSNLLLVKQKQAAVESGNGWAKDHPKSKAK 463
417 CYSCEGDGHRVQCINPSNLLLVKQKQAAVESGNGWAKDHPKSKAK 463
```

RESULT 4

```
Q8NET3 PRELIMINARY; PRT; 448 AA.
AC Q8NET3
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE BJ-HCC-25 tumor antigen.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Xueyuan D., Weifeng C.;
RT "Cloning and identification of genes which are differentially
RT expressed in carcinoma."
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY121806; AAM82754.1; --
SQ SEQUENCE 448 AA; 49334 MW; 5F6BC8A15B457A89 CRC64;
```

```
Query Match 36.5%; Score 883.5; DB 4; Length 448;
Best Local Similarity 42.6%; Pred. No. 3.7e-62;
Matches 201; Conservative 75; Mismatches 155; Indels 41; Gaps 10;

QY 1 MPTLLQDWCRCGEHLNTRCMLILGIPEDCGEDEFETLQACRHLGRYVIGRMFRREE 60
DB 1 MALTLLDWCCKGMDPRKALLIVGIPMECSEVEIQDTVKAGLQPLCAYRVLGRMFRRED 60

61 NAOAILLEAQDIDYALLPREIPGKGGPWVIVKPRNSDGEFLNRLNRLFLBEERTVSDM 120
DB 61 NAKAVFIELADTVNTTLPSHIPGKGGSWVIVKPRNPDDDFLSRLNLYFLKDEGRSMTDV 120

121 NRVLGSDTNCAPRVITISPEFTWAQTLGAAVQPLLEQMLYRELRVFSGNTISIPGALAF 180
DB 121 NRVLGSDTNCAPRVITISPEFTWAQTLGAAVQPLLEQMLYRELRVFSGNTISIPGALAF 180

121 ARALGC--CSLPAESLDAE--VMPQVRSPPLEPPKESMWYRKLKVFSGTASPSGGETF 175
DB 121 ARALGC--CSLPAESLDAE--VMPQVRSPPLEPPKESMWYRKLKVFSGTASPSGGETF 175

181 DAWLEHTTEMLQMWQVPEGEKRRRLMECLRGPAQVVGSLRAGNASITVECLAAQOVF 240
DB 176 EDWLQVTEIMPIWQVSEVEKRRRLLESIRGPAISIMRVLQANNDISITVECLDALKQIF 235

241 GPVESHKIAQVKLCAYQAEAGKVSFVLRLPQLQRAVENNVSRNNVOTRLKRVLSG 300
DB 236 GKDVEDFRASQFRFLQTSPIKGEKVSTFLRLPQLQRAVHKSPVSRVSTMDIRLKHLLAR 295

301 ATLPLDKLRDLKLMKQRRKPPGFLALVKLLR-BEEWEATLGPDRSLEGLVAPRPPAR 356
DB 296 VAMTPALRGKLELLDQRCPPNFLEMLKIRDEEWEATEAVMKNEKPSGRGRGASGRQ 355

357 PARITGVGAVPLPASGNSF-DARPS--QG-----YRRRGRGQHRGGVARGSGRSRKR 408
```

```
356 ARAEASVAPQATVQARSFSDSPQTIOGGLPPLVKRR-----LLGSESTR-- 402
409 KRHTFCYSCGDGHRVQCINPSNLLLVKQKQAAVES-GNGWAKDHPK 459
403 -----GED-HGQATYPKAENQTPGREGPQAAGEELGNEAGAGAMSHPK 444

RESULT 5
Q96PV4 PRELIMINARY; PRT; 452 AA.
AC Q96PV4
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein KIAA1934 (Fragment).
GN KIAA1934.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21456161; PubMed=11572484;
RT "Prediction of the coding sequences of unidentified human genes. XXI.
RT The complete sequences of 60 new cDNA clones from brain which code for
RT large proteins."
RL DNA Res. 8:179-187(2001).
DR EMBL: AB067521; BAB67827.1; --
DR Genbank: HGNC:18743; PNMA5.
KW Hypothetical protein.
FT NON TER
SQ SEQUENCE 452 AA; 50289 MW; A7553ECE7CA751EA CRC64;
```

```
Query Match 36.5%; Score 883.5; DB 4; Length 452;
Best Local Similarity 42.6%; Pred. No. 3.7e-62;
Matches 201; Conservative 75; Mismatches 155; Indels 41; Gaps 10;

QY 1 MPTLLQDWCRCGEHLNTRCMLILGIPEDCGEDEFETLQACRHLGRYVIGRMFRREE 60
DB 5 MALTLLDWCCKGMDPRKALLIVGIPMECSEVEIQDTVKAGLQPLCAYRVLGRMFRRED 64

61 NAOAILLEAQDIDYALLPREIPGKGGPWVIVKPRNSDGEFLNRLNRLFLBEERTVSDM 120
DB 65 NAKAVFIELADTVNTTLPSHIPGKGGSWVIVKPRNPDDDFLSRLNLYFLKDEGRSMTDV 124

121 NRVLGSDTNCAPRVITISPEFTWAQTLGAAVQPLLEQMLYRELRVFSGNTISIPGALAF 180
DB 125 ARALGC--CSLPAESLDAE--VMPQVRSPPLEPPKESMWYRKLKVFSGTASPSGGETF 179

181 DAWLEHTTEMLQMWQVPEGEKRRRLMECLRGPAQVVGSLRAGNASITVECLAAQOVF 240
DB 180 EDWLQVTEIMPIWQVSEVEKRRRLLESIRGPAISIMRVLQANNDISITVECLDALKQIF 239

241 GPVESHKIAQVKLCAYQAEAGKVSFVLRLPQLQRAVENNVSRNNVOTRLKRVLSG 300
DB 240 GKDVEDFRASQFRFLQTSPIKGEKVSTFLRLPQLQRAVHKSPVSRVSTMDIRLKHLLAR 299

301 ATLPLDKLRDLKLMKQRRKPPGFLALVKLLR-BEEWEATLGPDRSLEGLVAPRPPAR 356
DB 300 VAMTPALRGKLELLDQRCPPNFLEMLKIRDEEWEATEAVMKNEKPSGRGRGASGRQ 359

357 PARITGVGAVPLPASGNSF-DARPS--QG-----YRRRGRGQHRGGVARGSGRSRKR 408
DB 360 ARAEASVAPQATVQARSFSDSPQTIOGGLPPLVKRR-----LLGSESTR-- 406

409 KRHTFCYSCGDGHRVQCINPSNLLLVKQKQAAVES-GNGWAKDHPK 459
DB 407 -----GED-HGQATYPKAENQTPGREGPQAAGEELGNEAGAGAMSHPK 448

RESULT 6
```

O94959 ID O94959 PRELIMINARY; PRT; 364 AA.
 AC O94959;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Hypothetical protein KIAA0883 (Paraneoplastic associated brain-testis-cancer antigen).
 GN KIAA0883
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=99156230; PubMed=10048485;
 RA Nagase T., Ishikawa K., Suyama M., Kikuno R., Hirose M., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XII. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro.";
 RL DNA Res. 5:355-364 (1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Richen J.G., Dalmay J., Wade D., Rosenfeld M.R.;
 RT "Characterization of a Brain-Testis-Cancer Antigen.";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB020690; BAA74906.1; -;
 DR EMBL; AF286487; AAG28165.1; -;
 DR Genew; HGNC:9159; PNMA2.
 DR InterPro; IPR005162; Retrotrans gag.
 DR Pfam; PF03732; Retrotrans_gag; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 364 AA; 41509 MW; 6E417AD96E3F0E93 CRC64;
 Query Match 34.5%; Score 836.5; DB 4; Length 364;
 Best Local Similarity 49.7%; Pred. No. 1.6e-58;
 Matches 174; Conservative 63; Mismatches 108; Indels 5; Gaps 3;
 QY 1 MPTLLQDCRCGEHLNTRCMLILGIPEDCGEDEFETLQACRHLGRVYVIGRMFRRE 60
 DB 1 MALALLEDCRIMSVDQKSLMVTGIPVDEEAEIOEVLQETLKSIGRYLLGKIFRQ 60
 QY 61 NQAAILLELAQDIDYALLPREIPGKGPWEIVKPRNSDGEFLNRLNRFLEERRVSDM 120
 DB 61 NANAVLELLELTDVSAIPSEVQGGVWKVFKTNQDTEFLERLNLFLKESGQIVSGM 120
 QY 121 NRVLGSDTNCSPRVTVISPEFWT--WAQTLGAAVQPLLEQMLYRELRFVSGNTISIPGAL 178
 DB 121 FRALGQEGVSPATVPCISPELLAHLGQAMAHAPQLL-PMRYKRLRVFSGSAVPAPEE 179
 QY 179 AFDAWLEHTTEMLQWQVPEGEKRRRLMECLRGALQVGVSGLRASNASITVEECLAALQ 238
 DB 180 SPEVLEQATEIVKWPVTEAEKRWLAESLGRPALDLMHIVQADNPSISVEECLFAFK 239
 QY 239 VFGPVESHKIAQVKLCAYQAGEKVSFVLRLPQLLQRAVENNVVSRNVTQLKRVL 298
 DB 240 VFGSLESRTAQVRVLYKYQEGEKVSAYVLETLRLRAVEKRAIPRIADQVRLEQVM 299
 QY 299 SGATLPDKLRDKLKMQRKPPGFALVKKLREBEWEATLGPDRSLE 348
 DB 300 AGATLNQMLWCLRLKQDQGGPPFPFLMKVIREEEESASF--ENESIE 347
 RESULT 7
 Q9GMU3 ID Q9GMU3 PRELIMINARY; PRT; 364 AA.
 AC Q9GMU3;
 DT 01-MAY-2001 (TrEMBLrel. 16, Created)
 DT 01-MAY-2001 (TrEMBLrel. 16, Last sequence update)
 DE Hypothetical protein.
 GN KIAA0883
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=99156230; PubMed=10048485;
 RA Nagase T., Ishikawa K., Suyama M., Kikuno R., Hirose M., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XII. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro.";
 RL DNA Res. 5:355-364 (1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Richen J.G., Dalmay J., Wade D., Rosenfeld M.R.;
 RT "Characterization of a Brain-Testis-Cancer Antigen.";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB020690; BAA74906.1; -;
 DR EMBL; AF286487; AAG28165.1; -;
 DR Genew; HGNC:9159; PNMA2.
 DR InterPro; IPR005162; Retrotrans gag.
 DR Pfam; PF03732; Retrotrans_gag; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 364 AA; 41509 MW; 6E417AD96E3F0E93 CRC64;
 Query Match 34.5%; Score 836.5; DB 4; Length 364;
 Best Local Similarity 49.7%; Pred. No. 1.6e-58;
 Matches 174; Conservative 63; Mismatches 108; Indels 5; Gaps 3;
 QY 1 MPTLLQDCRCGEHLNTRCMLILGIPEDCGEDEFETLQACRHLGRVYVIGRMFRRE 60
 DB 1 MALALLEDCRIMSVDQKSLMVTGIPVDEEAEIOEVLQETLKSIGRYLLGKIFRQ 60
 QY 61 NQAAILLELAQDIDYALLPREIPGKGPWEIVKPRNSDGEFLNRLNRFLEERRVSDM 120
 DB 61 NANAVLELLELTDVSAIPSEVQGGVWKVFKTNQDTEFLERLNLFLKESGQIVSGM 120
 QY 121 NRVLGSDTNCSPRVTVISPEFWT--WAQTLGAAVQPLLEQMLYRELRFVSGNTISIPGAL 178
 DB 121 FRALGQEGVSPATVPCISPELLAHLGQAMAHAPQLL-PMRYKRLRVFSGSAVPAPEE 179
 QY 179 AFDAWLEHTTEMLQWQVPEGEKRRRLMECLRGALQVGVSGLRASNASITVEECLAALQ 238
 DB 180 SPEVLEQATEIVKWPVTEAEKRWLAESLGRPALDLMHIVQADNPSISVEECLFAFK 239
 QY 239 VFGPVESHKIAQVKLCAYQAGEKVSFVLRLPQLLQRAVENNVVSRNVTQLKRVL 298
 DB 240 VFGSLESRTAQVRVLYKYQEGEKVSAYVLETLRLRAVEKRAIPRIADQVRLEQVM 299
 QY 299 SGATLPDKLRDKLKMQRKPPGFALVKKLREBEWEATLGPDRSLE 348
 DB 300 AGATLNQMLWCLRLKQDQGGPPFPFLMKVIREEEESASF--ENESIE 347

OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
 OC Cercopitheidae; Macaca.
 OC NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Aizawa K., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Saito T., Okazaki Y., Nishi K., Kiyosawa H., Kondo S., Yanaka I., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 Query Match 34.1%; Score 826.5; DB 6; Length 364;
 Best Local Similarity 49.1%; Pred. No. 9.8e-58;
 Matches 172; Conservative 63; Mismatches 110; Indels 5; Gaps 3;
 QY 1 MPTLLQDCRCGEHLNTRCMLILGIPEDCGEDEFETLQACRHLGRVYVIGRMFRRE 60
 DB 1 MALALLEDCRIMSVDQKSLMVTGIPVDEEAEIOEVLQETLKSIGRYLLGKIFRQ 60
 QY 61 NQAAILLELAQDIDYALLPREIPGKGPWEIVKPRNSDGEFLNRLNRFLEERRVSDM 120
 DB 61 NANAVLELLELTDVSAIPSEVQGGVWKVFKTNQDTEFLERLNLFLKESGQIVSGM 120
 QY 121 NRVLGSDTNCSPRVTVISPEFWT--WAQTLGAAVQPLLEQMLYRELRFVSGNTISIPGAL 178
 DB 121 FRALGQEGVSPATVPCISPELLAHLGQAMAHAPQLL-PMRYKRLRVFSGSAVPAPEE 179
 QY 179 AFDAWLEHTTEMLQWQVPEGEKRRRLMECLRGALQVGVSGLRASNASITVEECLAALQ 238
 DB 180 SPEVLEQATEIVKWPVTEAEKRWLAESLGRPALDLMHIVQADNPSISVEECLFAFK 239
 QY 239 VFGPVESHKIAQVKLCAYQAGEKVSFVLRLPQLLQRAVENNVVSRNVTQLKRVL 298
 DB 240 VFGSLESRTAQVRVLYKYQEGEKVSAYVLETLRLRAVEKRAIPRIADQVRLEQVM 299
 QY 299 SGATLPDKLRDKLKMQRKPPGFALVKKLREBEWEATLGPDRSLE 348
 DB 300 AGATLNQMLWCLRLKQDQGGPPFPFLMKVIREEEESASF--ENESIE 347
 RESULT 8
 Q9CYP2 ID Q9CYP2 PRELIMINARY; PRT; 353 AA.
 AC Q9CYP2;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DE 5730402C15Rik protein.
 GN PNMAL OR 5730402C15R1K.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Embryo;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Aizawa K., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Saito T., Okazaki Y., Nishi K., Kiyosawa H., Kondo S., Yanaka I., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz K., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection";
RL Nature 409:685-690(2001);
DR EMBL; AK017476; BAB30762.1; --
DR MGD; MGI:2180564; Pnmal.
SQ SEQUENCE 353 AA; 39718 MW; F7B27378B8469675 CRC64;

Query Match 34.0%; Score 824; DB 11; Length 353;
Best Local Similarity 49.4%; Pred. No. 1.5e-57;
Matches 175; Conservative 55; Mismatches 116; Indels 8; Gaps 4;

QY 1 MPTLLQDWCRCGEHLNTRCMLILGIPDCGDEFEETLQACRHLGRYVRIGRMFRREE 60
DB 1 MAMTLLDWCRCGMDVNSQRALLVWGIPVNCDETEIETLQAMPQVS-YRVLGRMFREE 59

QY 61 NQAAILLELAQDIDYALLPREIPKGGPWEVIVKPRNSDGEFLNRLNRFLEERTVSDM 120
DB 60 NAKAALDELGTGVDYSLIPREMPKGLWKVVKFPTSDAFLRLHLFLAREGWTQDV 119

QY 121 NRVLGSDTNC SAPRVTISPEFWTWAQTLGAAVQPLLEQMLYRELRFVSGNTISIPGALAF 180
DB 120 ARVLGFQNPAPAPGPEPAEMLNY--ILDNVIQPLVESIWYKLTFLSGRDPGGETF 177

QY 181 DAWLEHTTEMLQMWQVPEGEKRRRLMECLRGPAQVSLGRASNASITVECLAAQOVF 240
DB 178 DSWLEHSNEIEEWMQVSDIEKRRRLMESLRGPAADVIRILKTNPAITTAECLEQV 237

QY 241 GPVESHKIAQVKLCKAYQAEKGVSSFVLRLEPLLQRAVENNVSRNNVTRLKRVLG 300
DB 238 GSVESSRDAQVRFLNTYQNPGEKLSYVIRLEPLLQKVDKGADKDNVQARLEQVIAG 297

QY 301 ATLPDKLRDLKLMKQRRKP-PGFLALVKLLRE---EWEATLGPDRSLEG 349
DB 298 ANHSGALRRQLWLTGATEGAPAPNFOLLVQVIREBEAKEEBAEAAALQLGLEG 351

RESULT 9
Q8VH24 PRELIMINARY; PRT; 353 AA.

ID Q8VH24
AC Q8VH24
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Paraneoplastic onconeural protein MAl.
GN MAl.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RL Pellkofer H.L., Schubart A., Pagany M., Voltz R.D., Linington C.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF335505; AAL73196.1; --
SQ SEQUENCE 353 AA; 39793 MW; FFC3A717FDCEDA3 CRC64;

Query Match 33.9%; Score 822; DB 11; Length 353;
Best Local Similarity 49.2%; Pred. No. 2.2e-57;
Matches 174; Conservative 55; Mismatches 117; Indels 8; Gaps 4;

QY 1 MPTLLQDWCRCGEHLNTRCMLILGIPDCGDEFEETLQACRHLGRYVRIGRMFRREE 60
DB 1 MAMTLLDWCRCGMDVNSQRALLVWGIPVNCDETEIETLQAMPQVS-YRVLGRMFREE 59

QY 61 NQAAILLELAQDIDYALLPREIPKGGPWEVIVKPRNSDGEFLNRLNRFLEERTVSDM 120
DB 60 NAKAALDELGTGVDYSLIPREMPKGLWKVVKFPTSDAFLRLHLFLAREGWTQDV 119

QY 121 NRVLGSDTNC SAPRVTISPEFWTWAQTLGAAVQPLLEQMLYRELRFVSGNTISIPGALAF 180
DB 120 ARVLGFQNPAPAPGPEPAEMLNY--ILDNVIQPLVESIWYKLTFLSGRDPGGETF 177

QY 181 DAWLEHTTEMLQMWQVPEGEKRRRLMECLRGPAQVSLGRASNASITVECLAAQOVF 240
DB 178 DSWLEHSNEIEEWMQVSDIEKRRRLMESLRGPAADVIRILKTNPAITTAECLEQV 237

QY 241 GPVESHKIAQVKLCKAYQAEKGVSSFVLRLEPLLQRAVENNVSRNNVTRLKRVLG 300
DB 238 GSVESSRDAQVRFLNTYQNPGEKLSYVIRLEPLLQKVDKGADKDNVQARLEQVIAG 297

QY 301 ATLPDKLRDLKLMKQRRKP-PGFLALVKLLRE---EWEATLGPDRSLEG 349
DB 298 ANHSGALRRQLWLTGATEGAPAPNFOLLVQVIREBEAKEEBAEAAALQLGLEG 351

RESULT 9
Q8VH24 PRELIMINARY; PRT; 353 AA.

ID Q8VH24
AC Q8VH24
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Paraneoplastic onconeural protein MAl.
GN MAl.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RL Pellkofer H.L., Schubart A., Pagany M., Voltz R.D., Linington C.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF335505; AAL73196.1; --
SQ SEQUENCE 353 AA; 39793 MW; FFC3A717FDCEDA3 CRC64;

Query Match 33.9%; Score 822; DB 11; Length 353;
Best Local Similarity 49.2%; Pred. No. 2.2e-57;
Matches 174; Conservative 55; Mismatches 117; Indels 8; Gaps 4;

QY 1 MPTLLQDWCRCGEHLNTRCMLILGIPDCGDEFEETLQACRHLGRYVRIGRMFRREE 60
DB 1 MAMTLLDWCRCGMDVNSQRALLVWGIPVNCDETEIETLQAMPQVS-YRVLGRMFREE 59

QY 61 NQAAILLELAQDIDYALLPREIPKGGPWEVIVKPRNSDGEFLNRLNRFLEERTVSDM 120
DB 60 NAKAALDELGTGVDYSLIPREMPKGLWKVVKFPTSDAFLRLHLFLAREGWTQDV 119

QY 121 NRVLGSDTNC SAPRVTISPEFWTWAQTLGAAVQPLLEQMLYRELRFVSGNTISIPGALAF 180
DB 120 ARVLGFQNPAPAPGPEPAEMLNY--ILDNVIQPLVESIWYKLTFLSGRDPGGETF 177

QY 181 DAWLEHTTEMLQMWQVPEGEKRRRLMECLRGPAQVSLGRASNASITVECLAAQOVF 240
DB 178 DSWLEHSNEIEEWMQVSDIEKRRRLMESLRGPAADVIRILKTNPAITTAECLEQV 237

QY 241 GPVESHKIAQVKLCKAYQAEKGVSSFVLRLEPLLQRAVENNVSRNNVTRLKRVLG 300
DB 238 GSVESSRDAQVRFLNTYQNPGEKLSYVIRLEPLLQKVDKGADKDNVQARLEQVIAG 297

QY 301 ATLPDKLRDLKLMKQRRKP-PGFLALVKLLRE---EWEATLGPDRSLEG 349
DB 298 ANHSGALRRQLWLTGATEGAPAPNFOLLVQVIREBEAKEEBAEAAALQLGLEG 351

RESULT 9
Q8VH24 PRELIMINARY; PRT; 353 AA.

ID Q8VH24
AC Q8VH24
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Paraneoplastic onconeural protein MAl.
GN MAl.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RL Pellkofer H.L., Schubart A., Pagany M., Voltz R.D., Linington C.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF335505; AAL73196.1; --
SQ SEQUENCE 353 AA; 39793 MW; FFC3A717FDCEDA3 CRC64;

Query Match 33.9%; Score 822; DB 11; Length 353;
Best Local Similarity 49.4%; Pred. No. 2.2e-57;
Matches 175; Conservative 54; Mismatches 117; Indels 8; Gaps 4;

QY 1 MPTLLQDWCRCGEHLNTRCMLILGIPDCGDEFEETLQACRHLGRYVRIGRMFRREE 60
DB 1 MAMTLLDWCRCGMDVNSQRALLVWGIPVNCDETEIETLQAMPQVS-YRVLGRMFREE 59

QY 61 NQAAILLELAQDIDYALLPREIPKGGPWEVIVKPRNSDGEFLNRLNRFLEERTVSDM 120
DB 60 NAKAALDELGTGVDYSLIPREMPKGLWKVVKFPTSDAFLRLHLFLAREGWTQDV 119

QY 121 NRVLGSDTNC SAPRVTISPEFWTWAQTLGAAVQPLLEQMLYRELRFVSGNTISIPGALAF 180
DB 120 ARVLGFQNPAPAPGPEPAEMLNY--ILDNVIQPLVESIWYKLTFLSGRDPGGETF 177

QY 181 DAWLEHTTEMLQMWQVPEGEKRRRLMECLRGPAQVSLGRASNASITVECLAAQOVF 240
DB 178 DSWLEHSNEIEEWMQVSDIEKRRRLMESLRGPAADVIRILKTNPAITTAECLEQV 237

QY 241 GPVESHKIAQVKLCKAYQAEKGVSSFVLRLEPLLQRAVENNVSRNNVTRLKRVLG 300
DB 238 GSVESSRDAQVRFLNTYQNPGEKLSYVIRLEPLLQKVDKGADKDNVQARLEQVIAG 297

QY 301 ATLPDKLRDLKLMKQRRKP-PGFLALVKLLRE---EWEATLGPDRSLEG 349
DB 298 ANHSGALRRQLWLTGATEGAPAPNFOLLVQVIREBEAKEEBAEAAALQLGLEG 351

RESULT 10
Q8C1C8 PRELIMINARY; PRT; 353 AA.

ID Q8C1C8
AC Q8C1C8
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Paraneoplastic onconeural protein MAl homolog.
GN PNMAL OR 5730402C15RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Brain;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK028331; BAC25885.1; --
DR MGD; MGI:2180564; Pnmal.
SQ SEQUENCE 353 AA; 39688 MW; 18CEDC3AC4E70939 CRC64;

Db 298 ANHSGALRRQLWLAGAEGPAPNLFQLLVQIRREEAKKEEAEAAALLQLGLEG 351

RESULT 11

Q8NG07 PRELIMINARY; PRT; 353 AA.
AC Q8NG07;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Paraneoplastic antigen.
GN PNMA1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Schueller M.M., Jenne D.E., Schutze dit Belkner N., Hohlfeild R.,
RT "Mal (PNMA1).";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Strausberg R.;
DT Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF320308; AA05100.1; -;
DR EMBL; BC039577; AA05100.1; -;
SQ SEQUENCE 353 AA; 39761 MW; EB7F5B6AEDA25961 CRC64;

Query Match 33.8%; Score 818; DB 4; Length 353;
Best Local Similarity 49.6%; Pred. No. 4.5e-57;
Matches 172; Conservative 52; Mismatches 113; Indels 10; Gaps 4;
QY 1 MPTLLQDRCRCHLNTNRCMLILGIPEDCGEDEFETQACRHILGRVIRVIGRMFRREE 60
DB 1 MAMTLLDRCRGMVNSORALLVWGIPVNCDEAEIETLQAAMPQVS-YRMLGRMFWEER 59
QY 61 NQAAILLEAQDIDYALLPREIPGKGPWEIVKPNRSDGEFLNRLNRLLEERRTVSDM 120
DB 60 NAKAALLELTGADVAAIPREMPGKGVWVLFKPTSDAEFLERLHLFLAREGWTVDV 119
QY 121 NRVLGSDTNCSPRVTISPEFTWATQTLGAAPVQLLEQMLRYRLRVFSGNTISIPGALAF 180
DB 120 ARVLGFQNPPTPGPMPAEMLY--ILDNVIQPLVESIWYKELTIFSGRDIPGGEETF 177
QY 181 DAWLEHTTLMQWQVPEGEKRRRLMECLRGALQVSGLRASNASITVEECIALAQVVF 240
DB 178 DPWLEHTNEVLBEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECIKALEQVF 237
QY 241 GPVESHKIAQVLCCKAYQAGKVSFVLRLPQLQRAVNNVSRNVTOTRLKRVLSG 300
DB 238 GSVESRDQAIKFLNTYQNGEKLQAYVIRLPLQKVKVKGAKIDKONVQARLEQVIAG 297
QY 301 ATLDPKRLDKLMKQRRKP-PGFLALVKLR-----EEBEWEATL 340
DB 298 ANHSGAIRQLWLTGAGEGAPNLFQLLVQIRREEAKKEEAEAAATL 344

RESULT 12

Q95144 PRELIMINARY; PRT; 353 AA.
AC Q95144;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Paraneoplastic neuronal antigen MAL.
GN MAL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;

QY 1 MPTLLQDRCRCHLNTNRCMLILGIPEDCGEDEFETQACRHILGRVIRVIGRMFRREE 60
DB 1 MAMTLLDRCRGMVNSORALLVWGIPVNCDEAEIETLQAAMPQVS-YRMLGRMFWEER 59
QY 61 NQAAILLEAQDIDYALLPREIPGKGPWEIVKPNRSDGEFLNRLNRLLEERRTVSDM 120
DB 60 NAKAALLELTGADVAAIPREMPGKGVWVLFKPTSDAEFLERLHLFLAREGWTVDV 119
QY 121 NRVLGSDTNCSPRVTISPEFTWATQTLGAAPVQLLEQMLRYRLRVFSGNTISIPGALAF 180
DB 120 ARVLGFQNPPTPGPMPAEMLY--ILDNVIQPLVESIWYKELTIFSGRDIPGGEETF 177
QY 181 DAWLEHTTLMQWQVPEGEKRRRLMECLRGALQVSGLRASNASITVEECIALAQVVF 240
DB 178 DPWLEHTNEVLBEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECIKALEQVF 237
QY 241 GPVESHKIAQVLCCKAYQAGKVSFVLRLPQLQRAVNNVSRNVTOTRLKRVLSG 300
DB 238 GSVESRDQAIKFLNTYQNGEKLQAYVIRLPLQKVKVKGAKIDKONVQARLEQVIAG 297
QY 301 ATLDPKRLDKLMKQRRKP-PGFLALVKLR-----EEBEWEATL 340
DB 298 ANHSGAIRQLWLTGAGEGAPNLFQLLVQIRREEAKKEEAEAAATL 344

Query Match 33.7%; Score 816; DB 4; Length 353;
Best Local Similarity 49.6%; Pred. No. 6.5e-57;
Matches 172; Conservative 52; Mismatches 113; Indels 10; Gaps 4;
QY 1 MPTLLQDRCRCHLNTNRCMLILGIPEDCGEDEFETQACRHILGRVIRVIGRMFRREE 60
DB 1 MAMTLLDRCRGMVNSORALLVWGIPVNCDEAEIETLQAAMPQVS-YRMLGRMFWEER 59
QY 61 NQAAILLEAQDIDYALLPREIPGKGPWEIVKPNRSDGEFLNRLNRLLEERRTVSDM 120
DB 60 NAKAALLELTGADVAAIPREMPGKGVWVLFKPTSDAEFLERLHLFLAREGWTVDV 119
QY 121 NRVLGSDTNCSPRVTISPEFTWATQTLGAAPVQLLEQMLRYRLRVFSGNTISIPGALAF 180
DB 120 ARVLGFQNPPTPGPMPAEMLY--ILDNVIQPLVESIWYKELTIFSGRDIPGGEETF 177
QY 181 DAWLEHTTLMQWQVPEGEKRRRLMECLRGALQVSGLRASNASITVEECIALAQVVF 240
DB 178 DPWLEHTNEVLBEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECIKALEQVF 237
QY 241 GPVESHKIAQVLCCKAYQAGKVSFVLRLPQLQRAVNNVSRNVTOTRLKRVLSG 300
DB 238 GSVESRDQAIKFLNTYQNGEKLQAYVIRLPLQKVKVKGAKIDKONVQARLEQVIAG 297
QY 301 ATLDPKRLDKLMKQRRKP-PGFLALVKLR-----EEBEWEATL 340
DB 298 ANHSGAIRQLWLTGAGEGAPNLFQLLVQIRREEAKKEEAEAAATL 344

RESULT 13

Q9BE35 PRELIMINARY; PRT; 150 AA.
AC Q9BE35;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cerebellum;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
libraries.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

```

RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Cerebellum;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
RT libraries.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR ENBL; AB060198; BAB41142.1; -.
DR EMBL; AB062932; BAB60727.1; -.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF00098; zf-CCHC; I.
DR PRINTS; PR00939; C2HCZNFINGER.
DR SMART; SM00343; Znf C2HC; 1.
DR PROSITE; PS00158; ZF_CCHC; 1.
KW Hypothetical protein_150
SQ SEQUENCE 150 AA; 16756 MW; D842F88B3C8A34A1 CRC64;

Query Match 32.7%; Score 793; DB 6; Length 150;
Best Local Similarity 98.0%; Pred. No. 1.3e-55;
Matches 147; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 314 MKQRKPPGFLALVKLIREEEWATIGPDRESLEGLVAPRPPARITGVGAVLPASGN 373
Db 1 MKQRKPPGFLALVKLIREEEWATIGPDRENLEGLVAPRPPARITGVGAVLPASGN 60

Qy 374 SFDARPSQVRRRGGRGHRGGVARAGSRGRKRKHTFCYSCGDGHIRVQCINPNSL 433
Db 61 SFDVRPSQVRRRGGRGHRGGVARAGSRGRKRKHTFCYSCGDGHIRVQCINPNSL 120

Qy 434 LLVQKQKQAAVESGNGNWDKSHPKSKAK 463
Db 121 LLVQKQKQAAVESGNGNWDKSHPKSKAK 150

RESULT 14
Q96BY2 ID Q96BY2 PRELIMINARY; PRT; 351 AA.
AC Q96BY2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Modulator of apoptosis 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR ENBL; BC015044; AAH15044.1; -.
DR Gene; HGNC:16658; MOAPL.
SQ SEQUENCE 351 AA; 39512 MW; 5310142AC02B563C CRC64;

Query Match 31.7%; Score 768.5; DB 4; Length 351;
Best Local Similarity 46.6%; Pred. No. 4e-53;
Matches 165; Conservative 60; Mismatches 118; Indels 11; Gaps 6;

Qy 1 MPTLLQDWCRCGEHLNTRCMLILGIPDCGDEFEETLQECRHLGRYRVIGRMFRREE 60
Db 1 MTLRLLEDWCRCGMDMNPRAKLLIAGISQSCSVAETEEALQAGLAPLGEYRLIGRMFRDE 60

Qy 61 NQAAILLELAQIDVALLPREIPGKGGPWEIVKPRNSDGEFLNRLNRFLEERETVSDM 120
Db 61 NRKVALVGLTARTSHALVPKEIPGKGIWRVIFKPPDPDNTFLSLNFEFLAGEGMTVGEL 120

Qy 121 NRVLGSDTNCSPRTVISPEFT--WAOTLGAAGVOPLEOMLYRELRFVSGNTGISPGAL 178
Db 121 SRALGHENGSLDPEQGMIPEMWAPMLAQAL-EALQPALQCLKYKLRVFSRESPEPGE 179

```

```

Qy 179 AFDWLHTTEMLQMWQVPEGEKRRRLMECLRGPAQVQVGLRASNASITVEECALAQ 238
Db 180 EFGRMWPHITQMIKAWQVDEKRRLLSLRGPALDVRVLKINPLIIVDECIQALEE 239

Qy 239 VFGPVESHKIAQVKLCKAYQAGEKVSFVLRLEPLQRAVENNVVRRNVNQTRLKRV 298
Db 240 VFGVTDNPRELVQKYLTYQDEKLSAYVLRLEPLQKLVQGAIBRDVAQNARLDQVI 299

Qy 299 SGATLPDKLRDKLKMQRKPPGFLALVKLIRE--EEWEATLGPDRSLEG 349
Db 300 AGA-VHKTIRRELNL-PEDGPAFGFLQLLVLIKDYAAAESEALL---QALLEG 348

RESULT 15
Q96A40 ID Q96A40 PRELIMINARY; PRT; 399 AA.
AC Q96A40;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ30760.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Ninomiya K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
RA Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamihara K.,
RA Katsuta N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Murakawa K.,
RA Kanehori K., Takahashi-Fujii A., Oshima H., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR ENBL; AK055322; BAB70902.1; -.
DR EMBL; BC007631; AAH07631.1; -.
KW Hypothetical protein.
SQ SEQUENCE 399 AA; 43875 MW; 256F5733C3EBB07D CRC64;

Query Match 31.7%; Score 768.5; DB 4; Length 399;
Best Local Similarity 42.5%; Pred. No. 4.9e-53;
Matches 171; Conservative 72; Mismatches 130; Indels 29; Gaps 7;

Qy 1 MPTLLQDWCRCGEHLNTRCMLILGIPDCGDEFEETLQECRHLGRYRVIGRMFRREE 60
Db 1 MAVTMLQDWCRCWMGVNARRGLLIGIPDCDABFQESLEAALRPMGHFTVLGKAFREED 60

Qy 61 NQAAILLELAQIDVALLPREIPGKGGPWEIVKPRNSDGEFLNRLNRFLEERETV 118
Db 61 NATAALVELDREYNVALVPRPIPGTGPWNVVFVPRCSGEBFLGLGRVHFHPEQGMWVE 120

Qy 119 DMNRVLGSDTNCSPRTVISPEFT--WAOTLGAAGVOPLEOMLYRELRFVSGNTGISPGAL 178
Db 121 SVAGALG-----VGLRRV-----CWLRSIQANQVQVAVRCQSLGVFSGRDQAPGEE 169

Qy 179 AFDWLHTTEMLQMWQVPEGEKRRRLMECLRGPAQVQVGLRASNASITVEECALAQ 237
Db 170 SFEVLDHTTEMLHVMQGVSRERRRLLEGLRGTLQVLHALLAENPARTAQDCLAALA 229

Qy 238 QVFGVESHKIAQVKLCKAYQAGEKVSFVLRLEPLQRAVENNVVRRNVNQTRLKRV 297
Db 230 QVFGDNEQATIRVKCLTAQQSGERLSFAFVLRLEVLQKAMEKALARASADRVLRRQM 289

Qy 298 LSGATLPDKLRDKLKMQRKPPGFLALVKLIREEEWATIG----PDRESLEGLEVA 353

```

Db 290 LTRAHLEPLDEALRKLRMAGRSPSPLEMLGLVRESEAWESLARSVRAQTQEGAGARAG 349
QY 354 PREPARI-TGVGAVP-----LPASGNSFDARPSQGYR 384
Db 350 AQAVARASTKVEAVPGGPGREPEGLLQAGGQAEAFELLQEGLK 391

Search completed: September 27, 2004, 17:10:11
Job time : 60.2533 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 27, 2004, 17:05:46 ; Search time 20.4265 Seconds
(without alignments)
2180.341 Million cell updates/sec

Title: US-10-037-860-13
Perfect score: 2423
Sequence: 1 MPTLLQDWCGRGHLNTRRC.....VESGNGNNAWDKSHPKSKAK 463

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	121	5.0	273	2 S00953	gag protein - frui
2	117	4.8	349	2 T18349	probable gag prote
3	117	4.8	1419	2 T32970	hypothetical prote
4	116	4.8	2287	2 T21312	hypothetical prote
5	114.5	4.7	474	2 S28419	lamin B-3 - mouse
6	114.5	4.7	661	2 T37753	hypothetical prote
7	113	4.7	476	2 A95982	probable GTP-bind
8	112.5	4.6	1356	2 F94486	probable retroelem
9	112	4.6	628	2 T23281	hypothetical prote
10	112	4.6	915	2 T26695	hypothetical prote
11	112	4.6	1041	2 T33699	hypothetical prote
12	111	4.6	1335	2 B84512	probable retroelem
13	110.5	4.6	838	2 G84599	probable retroelem
14	110	4.5	373	2 A44478	probable cell grow
15	109	4.5	406	2 E73366	conserved hypothet
16	109	4.5	1272	2 T49313	copia-type reverse
17	109	4.5	1320	2 F96614	probable copia-typ
18	109	4.5	1352	2 F86246	hypothetical prote
19	109	4.5	1352	2 T47925	copia-type polypro
20	109	4.5	2261	2 T20978	hypothetical prote
21	107	4.4	2241	2 T20971	hypothetical prote
22	106.5	4.4	857	2 S33821	median body protei
23	105.5	4.4	813	2 S36018	gene trantrack-p88
24	105.5	4.4	1864	2 F86378	hypothetical prote
25	105	4.3	759	2 T00153	protein F21J9.12 [
26	105	4.3	783	1 A38637	ras interactor RIN
27	105	4.3	1260	2 H89984	hypothetical prote
28	105	4.3	1291	2 S02021	micropia polyprote
29	104.5	4.3	627	2 B84482	probable gag-prote

30 104.5 4.3 864 2 B90395 purine NTPase (imp
31 104 4.3 483 2 AH3445 GTP-binding protei
32 104 4.3 1508 2 E87696 glutamate synthase
33 103 4.3 619 2 E82141 peptidyl-prolyl ci
34 102.5 4.2 4367 1 B54802 dynein heavy chain
35 102 4.2 200 2 T52627 splicing factor RS
36 102 4.2 551 2 G95099 DNA polymerase III
37 101.5 4.2 754 2 H84710 Mutator-like trans
38 101.5 4.2 1050 2 JW0092 serine-threonine k
39 101.5 4.2 1733 1 B45344 probable nuclear a
40 101 4.2 550 2 A46419 trophoblast-endoth
41 100.5 4.1 761 2 H65083 glycolate oxidase
42 100.5 4.1 1008 2 H85055 probable transposo
43 100 4.1 227 2 G71491 probable hth trans
44 100 4.1 294 2 D83136 hypothetical prote
45 100 4.1 859 2 T35785 probable beta-gluc

ALIGNMENTS

RESULT 1

S00953
gag protein - fruit fly (Drosophila melanogaster) transposon 1731
C;Species: Drosophila melanogaster
C;Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 21-Jul-2000
C;Accession: S00953
R;Fourcade-Peronnet, F.; d'Auriol, L.; Becker, J.; Galibert, F.; Best-Belpomme, M.
Nucleic Acids Res. 16, 6113-6125, 1988
A;Title: Primary structure and functional organization of Drosophila 1731 retrotransposor
A;Reference number: S00953; MUID:88289356; PMID:2456522
A;Accession: S00953
A;Molecule type: DNA
A;Residues: 1-273 <FOU>
A;Cross-references: EMBL:X07656; NID:g8700; PIDN:CAA30502.1; PID:g8701
C;Genetics:
A;Gene: gag
A;Cross-references: FlyBase:FBgn0000007

Query Match 5.0%; Score 121; DB 2; Length 273;
Best Local Similarity 20.8%; Pred. No. 0.25;
Matches 64; Conservative 52; Mismatches 102; Indels 90; Gaps 15;

Qy 179 AFDALHTTTLQ---MWQVPECEKRRRLMECLRGPALQ-----VVSGLRAS--- 223
Db 14 SYETWSIQMRSVLVHACLWKVVSQSVK--PEVDTGGAWOSQDEKALATITLSVKSSQLG 71
Qy 224 --NASITVECLALQVVF--GPVESHKIAQVKLCRAYQAGEKVSFVLRLEPLLQR- 277
Db 72 YVKGCLTAARAKVLDVHQPKGLRTVMLYKKLLSKELLE-GQSISHIKEFEKIFDAL 130
Qy 278 -AVENNVVSRNNVQTLKRLVSGATLPDKLRDKLKMQRKPPGFLAL-VKLLREEE 335
Db 131 DAVEIGITE-----XLRSVLLSLSPESFENFWAETRDDVPLFDALCIKIEEDTR 183
Qy 336 WEATLGPDRSLEGLVA-----PPPARITGVGNVPLPASGNSFDAPSPGQYRRR 387
Db 184 ---RGGAQQREKQTESAKAFTAVHKPQAPAR-----EARP----- 216
Qy 388 GRGQRRGGVARAGSRGRKRHTFCVSCEDGHIRVQC-----INPSNLLLVKKQKQAA 443
Db 217 -----SAKKRDKVVCYNGERHFRKANCERKVNKESATQEQCSLLNA 259
Qy 444 VESGNGNW 451
Db 260 LDSG-GFW 266

RESULT 2

T18349
probable gag protein - rice blast fungus gypsy retroelement
C;Species: Magnaporthe grisea (rice blast fungus)
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000

[illegible][illegible]


```

Qy 378 RPSQGVRRRGRGQ 391
Db 291 -PAKEHEILGEQ 303

RESULT 5
S28419
lamin B-3 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 20-Jun-2000
C;Accession: S28419
R;Furukawa, K.; Hotta, Y.
EMBO J. 12, 97-106, 1993
A;Title: cDNA cloning of a germ cell specific lamin B(3) from mouse spermatocytes and an
A;Reference number: S28419; MUID:93154351; PMID:8094052
A;Accession: S28419
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-474 <FUR>
A;Cross-references: EMBL:D13455; NID:g220471; PIDN:BAA02708.1; PID:g220472
C;Superfamily: cytoskeletal keratin

Query Match 4.7%; Score 114.5; DB 2; Length 474;
Best Local Similarity 23.9%; Pred. No. 1.5;
Matches 86; Conservative 48; Mismatches 147; Indels 79; Gaps 18;

Qy 145 AQTILGAAVQPLLEQMLYRELRFV-----SGNT-----ISIPGALAFD-AWLEHTTTEML 191
Db 22 ARPLMETVEGALPELGRPLREYVYKRGPGKTPVEDPVKSEGAAGVYPTWNNHLRVPT 81

Qy 192 QMVOVPEGEKR--RRMECLRGALQVVGSLGRASNASITVECLAL-----QQV-RGPV 243
Db 82 REQEVRETRRRHRLVE-----VDSRQOEYDFKMAQALDLRSQHDQVRLRYR 132

Qy 244 ESHKIAOVKL-----CKAQEAGEKVSFVLRLPL-----LOR---AVENNVV 284
Db 133 ELEQTYQAKLUNAKLSDNDKAHAREELKEARMRVESLYQLLGKQKQASAEHNI- 191

Qy 285 SRRNVNQTRLKRVLSGATLPDKLRDKLKMQRKPPQGLALVLLKLBREEEWEATLQPD- 343
Db 192 -----HELEALAGER--DKFRKWLDAKEQTEVRDAMQQLAEYOELLDIKALDM 242

Qy 344 -----RESLEG-----LEVAPPPARITGVAVPLPASGNSFDARPSQY-RRRRGRGQHR 393
Db 243 BISAYRKLLEGEERLKLSPSPSRITTSRATSSSSSGVGMVSGQGRKRRLTETD 302

Qy 394 RGGVARAG--SRGSRKRRHTFCYSGEDGHRVQCINPNSLLLVKQKQQAQAVESGNGW 451
Db 303 SGSPSRASRVSGSRLAQQT-----ATGVVNIDEVDPEG-RFVRLKNSSDKQOSLGNW 355

RESULT 6
T37753
hypochemical protein SPAC1687.10 - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T37753
R;Wedler, H.; Wambutt, R.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, December 1998
A;Reference number: Z21744
A;Accession: T37753
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-661 <WED>
A;Cross-references: EMBL:AL035064; PIDN:CAA22604.1; GSPDB:GN00066; SPDB:SPAC1687.10
A;Experimental source: strain 972h-; cosmid c1687
C;Genetics:
A;Gene: SPDB:SPAC1687.10
A;Map position: 1
A;Introns: 248/3

Query Match 4.7%; Score 114.5; DB 2; Length 661;
Best Local Similarity 21.5%; Pred. No. 2.3;

```

```

Matches 68; Conservative 35; Mismatches 131; Indels 83; Gaps 11;

Qy 100 GFLNRLNRFLEERRTVSDMR--VLGSDTNCAPRTVISPEFWTWAQTLGAQVQPLLE 157
Db 324 GELNNMIRRYLLESRVYKESQRMKVL-----GSASQPFLL 360

Qy 158 QMLYRELRFVSGNTISIPGALAFDAMLEHTTEMLQMOVPEGEKRRRLMECLRGPALQV 217
Db 361 KRLQKDFQLKAKLI-----CALREWEDNEKKIYVFGKPLSS---RLLMHSPILQNO 412

Qy 218 SGLRASNASITVECLAAQVFGPVEESHKIAQVVKLCKAYQAGEKVSFVRL-----EP 273
Db 413 ENISSNDNSKSSPESPAPKTKTKIENKLRNIDV-----QSKKIRFPRLRLASHEEP 466

Qy 274 LLQRAVENNVVSRNVNQTLLKRVLSGATLPDKLRDKLKMQRKPPGFLALVKLREE 333
Db 467 LIDRMVNQSPDTRSVSKSLVRETSSGCL-FKNHKSNTLSPPRKG-----SRH 515

Qy 334 EWEATLQPD-----RESLEGLEVAPRPARITGVGAVP----- 367
Db 516 GPREPCUSPDASSSIPVTDIKEILLPSQHDTPHNSVKLTGSSITTPASVLSRQMIPLLSR 575

Qy 368 LPASG---NSFDRPSQ 381
Db 576 TPPKGFTNSLDDTPTQ 592

RESULT 7
A95982
probable GNP-binding protein [imported] - Sinorhizobium meliloti (strain 1021) magaplasmi
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C;Accession: A95982
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan-
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo-
A;Reference number: A95982; MUID:21396508; PMID:11481431
A;Accession: A95982
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-476 <KUR>
A;Cross-references: GB:AL591985; PIDN:CAC49521.1; PID:g15141008; GSPDB:GN00167
A;Experimental source: strain 1021, megaplasmid pSymB
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: engA; Smb20995
A;Genome: Plasmid
C;Superfamily: Mycobacterium leprae probable GTP-binding protein; translation elongation

Query Match 4.7%; Score 113; DB 2; Length 476;
Best Local Similarity 23.7%; Pred. No. 1.9;
Matches 88; Conservative 41; Mismatches 101; Indels 142; Gaps 24;

Qy 80 REIPKGG-----GPWEIVKPRNSDGEFLNRLNRFLEERRTVSDMRVLGS 126
Db 183 RPAVEGTEDEBEVPAYDETPLRVAIVGRPNACK-STLINRFLGEDR-----LIT 232

Qy 127 DTNCSAPRTVISPEFWTWA-----OTLG-----AAVQPLLEQM-----LYRELRFVSGNT 171
Db 233 GPEAGITRDSISVE-WDWRGRTIKMFDTAGRRKAKVQEKLEKLSVADALRAIR-FASTV 290

Qy 172 I-----SIP-----GALAFDAWLEHTTEMLQWQ---VPEGEKRRRL 205
Db 291 VIVFDATIPFEKQDQLIVDLVLREGRAAVLAFNKW-----DLVENWQALLDLREKTERL 345

```


Db 310 KENVESFERVEKLV--AMTSGKSIDYIQTAHTFLEG-----LGDHIQWEVKAK 361
 QY 319 KPGFLALVKLLREEEWEATLQPD-----RESLEGLVAPRPAPRITGVGAVPLPAS 371
 Db 362 RP-----LTLEAYDALSOQLLODKMRKQIED-----PVA----- 393
 QY 372 GNSFDARPSGQYRRRRGRGQHRRGVAVAGRSRKRKRHTFCYSCGEDGHRVQCINPS 431
 Db 394 -NAFFAR-----HGLQKNSFVGNVCYCGKRGHTANEC----- 424
 QY 432 NLLLVKQKQAAVESGNG 449
 Db 425 -----RRKKSDESQNG 435

RESULT 11
 T33699
 hypothetical protein F49F1.8 - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C;Accession: T33699
 R;Miller, N.; Wamsley, P.
 submitted to the EMBL Data Library, October 1998
 A;Description: The sequence of C. elegans cosmid F49F1.
 A;Reference number: Z21389
 A;Accession: T33699
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-1041<ML>
 A;Cross-references: EMBL:AF100656; PIDN:AAC68951.1; GSPDB:GN00022; CESP:F49F1.8
 A;Experimental source: strain Bristol N2; clone F49F1
 C;Genetics:
 A;Gene: CESP:F49F1.8
 A;Map position: 4
 A;Introns: 172/3; 293/1; 334/3; 579/1

Query Match 4.6%; Score 112; DB 2; Length 1041;
 Best Local Similarity 22.0%; Pred. No. 6.5;
 Matches 56; Conservative 50; Mismatches 88; Indels 60; Gaps 12;

QY 84 GKGGPWEIVKPRNSDGEFLNRLNRFLEERR-----TVSP----- 119
 Db 777 GEGGELDAVARVTQSKVLETRDILERSNVASATPPSTASDGAGSTFEENVIOQVAQ 836
 QY 120 -----MNRVLG-----SDTNCAPRTVISPEFTWTAQTLGAARV-----QPLLEQWLY 161
 Db 837 ANQGSSEMGQSVSENSDSDTNAAPLSVAVDKK---RSGSVGNAKSKSYEDRPILGMLRL 893
 QY 162 R---ELRVFGNTISIPCALAFDAWLEHTTMLQWQ-VPEGEKRRRLMECLRGPAQV 217
 Db 894 RPAVOLTPYKGTM-----AEDFSFIRSFKDQCEASEKLTDDQLKFLTLCLLDRARNA 948
 QY 218 SGLRASNAITVECLAAQOVF-GPV-ESHKIAQVKCKAYQAGEKVSFVLRLPL 275
 Db 949 EDETKLNAVTLQILAKRMTTFENPVLIKHNQLRCK--QXPERSVEAFHRIEEL 1006
 QY 276 QRA--VENNVVSR 287
 Db 1007 RAASTVENIVTKR 1020

RESULT 12
 B84512
 Probable retroelement pol polyprotein [imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
 C;Accession: B84512
 R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A;Reference number: A84420; MUID:20083487; PMID:10617197
 A;Accession: B84512
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-1335<STO>
 A;Cross-references: GB:AE002093; NID:g4388818; PIDN:AAD19773.1; GSPDB:GN00139
 C;Genetics:
 A;Gene: At2g13930
 A;Map position: 2
 C;Superfamily: retrovirus-related polyprotein

Query Match 4.6%; Score 111; DB 2; Length 1335;
 Best Local Similarity 18.8%; Pred. No. 11;
 Matches 59; Conservative 42; Mismatches 104; Indels 108; Gaps 11;

QY 197 PEGEKRRRLMECLRGPAQVVSGLRASNASITVECLAAQOVFGPVPVESHKIAQVKCKA 256
 Db 29 PEKRRKRDADAVAR-----LERCDKAKNVIPLNVADKVLKRIELCKT 70
 QY 257 YQAGEKVSFVRLRP-----LLQRAVENNV-----VSRNVNQT 292
 Db 71 AAFAWETLDFMIRSLPHRVYTQLSFYTFKMQENKKIDENIDFLKIVADLAHLQIDVT 130
 QY 293 -RLKRVLSGATLP---DKLRDKLKMQRKPPGFLALVKLLREEEWEATLGPDRSLE 348
 Db 131 DEVOAILLLSSLPARYDGLVETWKYSNRK-----LRLDDVMVAARDKERE--- 177
 QY 349 GLEVAPRPAPRITGVGAVPLPAGNSFDARPSQYRRRRGRGQHRRGVAVAGRSRKR 408
 Db 178 -LSQNNRP-----VVEGHFARGRPDGNKNNQGNKGRSKSADGRK----- 219
 QY 409 KRHTFCYSCGEDGHRVQC-----INPSNLLLVKQKQ 441
 Db 220 -----VCWICGEGHFKQCYKWIERNKSKQGSNGESSLAKSTEAFNPAMVLLATDTL 275
 QY 442 AAVERSGNGNNAWD 454
 Db 276 VVTDSIANEWLD 288

RESULT 13

G84599
 Probable retroelement pol polyprotein [imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C;Accession: G84599
 R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A;Reference number: A84420; MUID:20083487; PMID:10617197
 A;Accession: G84599
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-838<STO>
 A;Cross-references: GB:AE002093; NID:g4567266; PIDN:AAD23679.1; GSPDB:GN00139
 C;Genetics:
 A;Gene: At2g1310
 A;Map position: 2

Query Match 4.6%; Score 110.5; DB 2; Length 838;
 Best Local Similarity 21.1%; Pred. No. 6.2;
 Matches 61; Conservative 49; Mismatches 120; Indels 59; Gaps 10;

QY 214 LQVVSGLRASNASITVECLAAQOVFGPVPVESHKIAQVKCKAYQAGEKVSFVLRL--- 270
 Db 31 LGLLEGLEDEEA-IEEESTETDTSILTKTED-KVLEKRCARSTVLSLGNHVLKVI 88
 QY 271 ---LEPILQRAVENNVSRNVNQTRLKRVLSG-----ATLPDKLRDKLKLKMKORR--- 318
 Db 89 KEKTAAGMIRVLDKLFMAKSLPNRIYKORLYGYKMSDSMTIEENVNDFFKLISLENVK 148

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 27, 2004, 17:10:17 ; Search time 134.663 Seconds
(without alignments)
1105.584 Million cell updates/sec

Title: US-10-037-860-13
Perfect score: 2423
Sequence: 1 MPTLLQDWCRGEHLNTRRC.....VESGNGNWDKSHPKSKAK 463

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1349238 seqs, 321558718 residues
Total number of hits satisfying chosen parameters: 1349238

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	2423	100.0	463	13	US-10-037-860-13
2	883.5	36.5	452	16	US-10-037-860-13
3	818	33.8	353	9	US-09-965-529-7
4	818	33.8	353	10	US-09-969-680A-7
5	768.5	31.7	351	9	US-09-965-529-1
6	768.5	31.7	351	10	US-09-969-680A-1
7	768.5	31.7	351	12	US-09-804-014A-16
8	768.5	31.7	351	15	US-10-341-434-10
9	768.5	31.7	399	15	US-10-094-749-1978
10	766.5	31.6	329	13	US-10-037-860-4
11	765.5	31.6	328	12	US-09-804-014A-40
12	742	30.6	321	12	US-09-804-014A-39
13	740.5	30.6	312	12	US-09-804-014A-73
14	740.5	30.6	312	12	US-09-804-014A-74
15	618.5	25.5	283	13	US-10-037-860-11

16	462.5	19.1	195	13	US-10-037-860-7	Sequence 7, Appli
17	423.5	17.5	403	15	US-10-094-466-38	Sequence 38, Appl
18	384.5	15.9	337	12	US-10-296-115-1208	Sequence 1208, Ap
19	332	13.7	120	12	US-09-804-014A-42	Sequence 42, Appl
20	324	13.4	120	12	US-09-804-014A-41	Sequence 41, Appl
21	323	13.3	149	13	US-10-037-860-9	Sequence 9, Appli
22	322	13.3	206	14	US-10-029-386-33747	Sequence 33747, A
23	270	11.1	116	9	US-09-864-761-34645	Sequence 34645, A
24	191	7.9	538	16	US-10-408-765A-2992	Sequence 2992, Ap
25	158.5	6.5	1332	15	US-10-374-780A-1037	Sequence 1037, Ap
26	156	6.4	1394	16	US-10-437-963-185722	Sequence 185722,
27	156	6.4	1433	15	US-10-374-780A-1040	Sequence 1040, Ap
28	156	6.4	1433	16	US-10-437-963-110685	Sequence 110685,
29	156	6.4	1828	16	US-10-437-963-123225	Sequence 123225,
30	155.5	6.4	1150	16	US-10-437-963-122585	Sequence 122585,
31	155	6.4	584	12	US-10-221-278-355	Sequence 355, App
32	155	6.4	584	15	US-10-291-172-355	Sequence 355, App
33	154	6.4	1360	16	US-10-437-963-185720	Sequence 185720,
34	154	6.4	2003	16	US-10-437-963-117298	Sequence 117298,
35	152	6.3	1433	15	US-10-374-780A-1035	Sequence 1035, Ap
36	152	6.3	1696	16	US-10-437-963-17296	Sequence 17296,
37	151	6.2	1636	16	US-10-437-963-122665	Sequence 122665,
38	151	6.2	1695	16	US-10-437-963-122632	Sequence 122632,
39	149	6.1	1711	16	US-10-437-963-199076	Sequence 199076,
40	149	6.1	3783	16	US-10-437-963-146318	Sequence 146318,
41	148	6.1	1777	16	US-10-437-963-122620	Sequence 122620,
42	148	6.1	2155	16	US-10-437-963-146341	Sequence 146341,
43	147	6.1	1236	16	US-10-437-963-122619	Sequence 122619,
44	147	6.1	2664	16	US-10-437-963-117303	Sequence 117303,
45	146.5	6.0	1008	16	US-10-437-963-184092	Sequence 184092,

ALIGNMENTS

RESULT 1
US-10-037-860-13
; Sequence 13, Application US/10037860
; Publication No. US20020123114A1
; GENERAL INFORMATION:
; APPLICANT: Jerome B. Posner
; APPLICANT: Josep O. Dalmau
; APPLICANT: Myrna R. Rosenfeld
; TITLE OF INVENTION: Ma FAMILY POLYPEPTIDES AND ANTI-Ma
; TITLE OF INVENTION: ANTIBODIES
; FILE REFERENCE: 2581.1004-004
; CURRENT APPLICATION NUMBER: US/10/037,860
; CURRENT FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 09/189,527
; PRIOR FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 463
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-037-860-13

Query Match	100.0%	Score 2423;	DB 13;	Length 463;
Best Local Similarity	100.0%;	Pred. No. 6.6e-220;		
Matches 463;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MPTLLQDWCRGEHLNTRRCMLILGIPEDCGDEFEETLQACHLGRYVIGMFREE	60	
Db	1	MPTLLQDWCRGEHLNTRRCMLILGIPEDCGDEFEETLQACHLGRYVIGMFREE	60	
Qy	61	NAQAIILELAQIDYALLPREIPGKGGPWEVIVPRNSDGEFLNLRNLFEEERTVSDM	120	
Db	61	NAQAIILELAQIDYALLPREIPGKGGPWEVIVPRNSDGEFLNLRNLFEEERTVSDM	120	
Qy	121	NRVLGSDTNCSPRVITISPEFWTWAQTIGAAVQPLLEQMLYRELRFVSGNTISIPGALAF	180	
Db	121	NRVLGSDTNCSPRVITISPEFWTWAQTIGAAVQPLLEQMLYRELRFVSGNTISIPGALAF	180	


```
; Sequence 7, Application US/09969680A
; Publication No. US20030124649A1
; GENERAL INFORMATION:
; APPLICANT: LAL, Preeti; YUE, Henry
; APPLICANT: TANG, Y. Tom; BANDMAN, Olga
; APPLICANT: BURFORD, Neil; AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.; LU, Dyung Aina M.
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0731-1 USA
; CURRENT APPLICATION NUMBER: US/09/969,680A
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US00/22315
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/149,641
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/164,203
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PERL Program
; SEQ ID NO 7
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030124649A1 2483172CD1
US-09-969-680A-7

Query Match      33.8%; Score 818; DB 10; Length 353;
Best Local Similarity 49.6%; Pred. No. 2.4e-68;
Matches 172; Conservative 52; Mismatches 113; Indels 10; Gaps 4;

Qy 1 MPTLLQDWCRGHEHLNTRRCMLILGIPEDCGEDEFEEFTLOEACRHGLGRYRVIGRMFRREE 60
Db 1 MAMTLLEDWCRGMDVNSQRAALLVWGIPVNCDEABIEETLQAAMPQVS-YRMLGRMFWRREE 59

Qy 61 NQAAILLELAQDIDYALLPREIPKGGPWEVIVKPRNSDGEFLNRLNRFLEEEERTVSDM 120
Db 60 NAKAALLELTGADVAAITPREMPKGGVWKVLFKPTSDAEFLERLHLFLAREGWTQDV 119

Qy 121 NRVLGSDTNC SAPRTVISPEFTWAQTLGAQVQLLEQMLYRELVRVSGNTISIPGALAF 180
Db 120 ARVLGFGNQPTTPGEMPAEMLN-Y--ILDNVIQPLVESIWYKRLTFLSGRDIPOGGETP 177

Qy 181 DAWLEHTTEMLQMWQVPEGEKRRRLMECLRGPAQVLSGLRASNASITVEECLAALQOVF 240
Db 178 DPWLEHTNEVLQWQVSDVEKRRRLMESLRGPAADVIRILKSNNPATITAECLKALEQVF 237

Qy 241 GPVESHKIAQVKLCKAYQAGEKVSFVLRLEPLQRAVENNVSRNNVQTRLRKRVLSG 300
Db 238 GSVSSRDAQIKFLNTYQNPGEKLSAYVIRLEPLQKVKVEKGAIDKONVQARLEQVIAG 297

Qy 301 ATLPDKLRDLKLMKQRRKP-PGFLALVKKLR-----EEWEATL 340
Db 298 ANHSGAIRRQLWLTGAGEGPAPNLFQLLVQIRREAEKEEAEATL 344

RESULT 5
US-09-965-529-1
; Sequence 1, Application US/09965529
; Publication No. US20020182671A1
; GENERAL INFORMATION:
; APPLICANT: LAL, Preeti
; APPLICANT: TANG, Y. Tom
; APPLICANT: BANDMAN, Olga
; APPLICANT: BURFORD, Neil
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS
```

```
; FILE REFERENCE: PF-0731 USA
; CURRENT APPLICATION NUMBER: US/09/965,529
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/149,641; 60/164,203; PCT/US00/22315
; PRIOR FILING DATE: 1999-08-17; 1999-11-09; 2000-08-14
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020182671A1 1123010CD1
US-09-965-529-1

Query Match      31.7%; Score 768.5; DB 9; Length 351;
Best Local Similarity 46.6%; Pred. No. 1.1e-63;
Matches 165; Conservative 60; Mismatches 118; Indels 11; Gaps 6;

Qy 1 MPTLLQDWCRGHEHLNTRRCMLILGIPEDCGEDEFEEFTLOEACRHGLGRYRVIGRMFRREE 60
Db 1 MTLRLLEDWCRGMDMNPRAKLLIAGISQSCVAIEEALQAGLAPLGEYRLLRGMRFRDE 60

Qy 61 NQAAILLELAQDIDYALLPREIPKGGPWEVIVKPRNSDGEFLNRLNRFLEEEERTVSDM 120
Db 61 NRKVALVGLTAETSHALVPEIKPGGGINRVIFKPPDPDNTFLSRLNEFLAGEGWTGEL 120

Qy 121 NRVLGSDTNC SAPRTVISPEFT--WAQTLGAQVQLLEQMLYRELVRVSGNTISIPGAL 178
Db 121 SRALGHENGSLDPEQGMIPENWAPMLAAL-EALQPALQCLKYKLRVFSGRSPPEGEE 179

Qy 179 AFDWLEHTTEMLQMWQVPEGEKRRRLMECLRGPAQVLSGLRASNASITVEECLAALQ 238
Db 180 EFGEMWFHTTQMIKAWQVDPVEKRRRLLESIRGPAADVIRLKLNNPLITVDECLQALEE 239

Qy 239 VFGVESHKIAQVKLCKAYQAGEKVSFVLRLEPLQRAVENNVSRNNVQTRLRKRVL 298
Db 240 VFGVTDPNRELQVKLTYYQKDEKLSAYVIRLEPLQKLVORGAIERDANQARLQVI 299

Qy 299 SGATLPDKLRDLKLMKQRRKP-PGFLALVKKLR--EEWEATLGPDRSLEG 349
Db 300 AGA-VHKTIRRELNL-PEDGPAPGFLQLLVLIKDYEAEEAEALL---QALLEG 348

RESULT 6
US-09-969-680A-1
; Sequence 1, Application US/09969680A
; Publication No. US20030124649A1
; GENERAL INFORMATION:
; APPLICANT: LAL, Preeti; YUE, Henry
; APPLICANT: TANG, Y. Tom; BANDMAN, Olga
; APPLICANT: BURFORD, Neil; AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.; LU, Dyung Aina M.
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0731-1 USA
; CURRENT APPLICATION NUMBER: US/09/969,680A
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US00/22315
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/149,641
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/164,203
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
```

OTHER INFORMATION: Incyte ID No. US20030124649A1 1123010C01
US-09-969-680A-1

Query Match 31.7%; Score 768.5; DB 10; Length 351;
Best Local Similarity 46.6%; Pred. No. 1.1e-63;
Matches 165; Conservative 60; Mismatches 118; Indels 11; Gaps 6;

```
QY 1 MPTLLQDWCRCGHEHNTNRCMLILGIPEDCGDEFEETLQEAACHLGRYRVIGRMFRREE 60
Db 1 MTRLLEDWCRCGMDMNPRAKLLIAGISQSCSVAIEEALQAGLAPLGEYRLDGRMFRDE 60

QY 61 NQAAILLELAQDIDYALLPREIPGKGPWEVIVKPRNSDGEFLNRLNRFLEEERRTVSDM 120
Db 61 NTKVALVGLTAETSHALVPKEIPGKGIWRVIFKPPDPNTPTFLSRLNEFLAGEMTVGL 120

QY 121 NRVLGSDTNCSPRVITISPEFWT--WAQTLGAAVQPLLEOMLYRELRFVSGNTISIPGAL 178
Db 121 SRALGHENGLDPEQGMIPEMWAPMLAQAAL-EALQPALQCLKYKLRVFSGRSPBPGE 179

QY 179 AFDWLEHTTEMLQMWQVPEGEKRRRLMECLRGPAQVVGSLRASNASITVEECALAOQ 238
Db 180 EFGRMWFHTTQMKAWQVPEGEKRRRLLESRLGPAQVIRVLLKINPLITVDECLQALEE 239

QY 239 VFGPVESHKIAQVKLCAYQEAQEKVSFVLRLLEPLLQRAVENNVSRNVTQRLKRVL 298
Db 240 VFGVTDNPRELQVKYLTYYQDEEKLAVVLRLEPLLQKLVQGAERDAVNAQRLDQVI 299

QY 299 SGATLPDKLRDKLKMQRKPPGFALVVKLRE---EEWEATLGPDRSLEG 349
Db 300 AGA-VHKTIRRELNL-PEDGPAPGFLQLLVLIKDYEAABEEBALL---QAILEG 348
```

RESULT 7

US-09-804-014A-16
; Sequence 16, Application US/09804014A
; Publication No. US20030064489A1
; GENERAL INFORMATION:
; APPLICANT: Li, Li
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Vernet, Corine
; APPLICANT: Ferrandes, Elma
; APPLICANT: Shimkets, Richard
; APPLICANT: Spaderina, Steven
; APPLICANT: Majumder, Kumud
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-721 US
; CURRENT APPLICATION NUMBER: US/09/804,014A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/188,316
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/188,277
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/189,139
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/189,140
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/190,401
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/190,231
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-804-014A-16

Query Match 31.7%; Score 768.5; DB 12; Length 351;
Best Local Similarity 46.6%; Pred. No. 1.1e-63;
Matches 165; Conservative 60; Mismatches 118; Indels 11; Gaps 6;

```
QY 1 MPTLLQDWCRCGHEHNTNRCMLILGIPEDCGDEFEETLQEAACHLGRYRVIGRMFRREE 60
```

```
Db 1 MTRLLEDWCRCGMDMNPRAKLLIAGISQSCSVAIEEALQAGLAPLGEYRLDGRMFRDE 60

QY 61 NQAAILLELAQDIDYALLPREIPGKGPWEVIVKPRNSDGEFLNRLNRFLEEERRTVSDM 120
Db 61 NTKVALVGLTAETSHALVPKEIPGKGIWRVIFKPPDPNTPTFLSRLNEFLAGEMTVGL 120

QY 121 NRVLGSDTNCSPRVITISPEFWT--WAQTLGAAVQPLLEOMLYRELRFVSGNTISIPGAL 178
Db 121 SRALGHENGLDPEQGMIPEMWAPMLAQAAL-EALQPALQCLKYKLRVFSGRSPBPGE 179

QY 179 AFDWLEHTTEMLQMWQVPEGEKRRRLMECLRGPAQVVGSLRASNASITVEECALAOQ 238
Db 180 EFGRMWFHTTQMKAWQVPEGEKRRRLLESRLGPAQVIRVLLKINPLITVDECLQALEE 239

QY 239 VFGPVESHKIAQVKLCAYQEAQEKVSFVLRLLEPLLQRAVENNVSRNVTQRLKRVL 298
Db 240 VFGVTDNPRELQVKYLTYYQDEEKLAVVLRLEPLLQKLVQGAERDAVNAQRLDQVI 299

QY 299 SGATLPDKLRDKLKMQRKPPGFALVVKLRE---EEWEATLGPDRSLEG 349
Db 300 AGA-VHKTIRRELNL-PEDGPAPGFLQLLVLIKDYEAABEEBALL---QAILEG 348
```

RESULT 8

US-10-341-434-10
; Sequence 10, Application US/10341434
; Publication No. US20030215835A1
; GENERAL INFORMATION:
; APPLICANT: OriGene Technologies
; TITLE OF INVENTION: Differentially Regulated Prostate Cancer Genes
; FILE REFERENCE: 9U 204 205 R1
; CURRENT APPLICATION NUMBER: US/10/341,434
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/348,164
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 60/348,119
; PRIOR FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 10
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-341-434-10

Query Match 31.7%; Score 768.5; DB 15; Length 351;
Best Local Similarity 46.6%; Pred. No. 1.1e-63;
Matches 165; Conservative 60; Mismatches 118; Indels 11; Gaps 6;

```
QY 1 MPTLLQDWCRCGHEHNTNRCMLILGIPEDCGDEFEETLQEAACHLGRYRVIGRMFRREE 60
Db 1 MTRLLEDWCRCGMDMNPRAKLLIAGISQSCSVAIEEALQAGLAPLGEYRLDGRMFRDE 60

QY 61 NQAAILLELAQDIDYALLPREIPGKGPWEVIVKPRNSDGEFLNRLNRFLEEERRTVSDM 120
Db 61 NTKVALVGLTAETSHALVPKEIPGKGIWRVIFKPPDPNTPTFLSRLNEFLAGEMTVGL 120

QY 121 NRVLGSDTNCSPRVITISPEFWT--WAQTLGAAVQPLLEOMLYRELRFVSGNTISIPGAL 178
Db 121 SRALGHENGLDPEQGMIPEMWAPMLAQAAL-EALQPALQCLKYKLRVFSGRSPBPGE 179

QY 179 AFDWLEHTTEMLQMWQVPEGEKRRRLMECLRGPAQVVGSLRASNASITVEECALAOQ 238
Db 180 EFGRMWFHTTQMKAWQVPEGEKRRRLLESRLGPAQVIRVLLKINPLITVDECLQALEE 239

QY 239 VFGPVESHKIAQVKLCAYQEAQEKVSFVLRLLEPLLQRAVENNVSRNVTQRLKRVL 298
Db 240 VFGVTDNPRELQVKYLTYYQDEEKLAVVLRLEPLLQKLVQGAERDAVNAQRLDQVI 299

QY 299 SGATLPDKLRDKLKMQRKPPGFALVVKLRE---EEWEATLGPDRSLEG 349
Db 300 AGA-VHKTIRRELNL-PEDGPAPGFLQLLVLIKDYEAABEEBALL---QAILEG 348
```

```
RESULT 9
US-10-094-749-1978
; Sequence 1978, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; PRIOR FILING DATE: 2002-03-12
; PRIOR FILING DATE: 2002-01-24
; PRIOR FILING DATE: 2001-09-14
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1978
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-1978

Query Match      31.7%; Score 768.5; DB 15; Length 399;
Best Local Similarity 42.5%; Pred. No. 1.4e-63;
Matches 171; Conservative 72; Mismatches 130; Indels 29; Gaps 7;

Qy      1  MPTLLQDWCRCGEHLNTRRCMLILGIPEDCCGDEFEETLQACRHLGRYRVIQGRMFRREE 60
Db      1  MAMTLLEDWCRCGMDVNSQRTLLVWGIPVNCDEAEIETLQAAMPQVS-YRMLGRMFWREE 59

Qy      61  NQAAILLELAQDIDYALLPREIPKGGPWEVIVKPRNSDGEF--LNRLNRFLEERRTVS 118
Db      61  NATAAVELDREVNVALVPREIPGTGGPNVVFVPRCSGEEFLGLGRVFFHFEQEQWVE 120

Qy      119  DMNRVLGSDTNCSPRVTTISPEFTWTAQTLGAAVQPLLEQMLYRELRFVSGNTISIPGAL 178
Db      121  SVAGALG-----VGLRRV-----CWLRSIQAVQVPWEAVRCQSLGVFSGRDQAPGEE 169

Qy      179  AFDAWLEHTTEMLQWQV-PEGEKRRRLMECLRGPAQVQVSGLRASNASITVBECLAALQ 237
Db      170  SFEVLDHTTEMLHWQVGSERERRRRLLEGLRTGTLQVHALLAENPARTAQDCLAALA 229

Qy      238  QVGEVESHKIAQVKLCAYQAEAGKVSFVLRLEPLLQRAVENNVVSRNNVQTRLKRV 297
Db      230  QVFGDNEQATIRVKCLTAQQQSGERLSAFVLRLEVLQKAMEKALARASADRVRLRQM 289

Qy      298  LSGATLPDKLRDKLKMQRKPPGFTALVKLLREEEWEATLG----PDRESLEGLEVA 353
Db      290  LTRAHLTEPLDEALRKLKRMAGRSFLEMIGLVRESEAWASLASRVAQTOEGAGARAG 349

Qy      354  PRPPARI--TGVAVP-----LPSAGNSFDRAPSGQYR 384
Db      350  AQAVARASTKVENVPGFGREPEGLLQAGGQAEELLQGLK 391
```

```
RESULT 10
US-10-037-860-4
; Sequence 4, Application US/10037860
; Publication No. US20020123114A1
; GENERAL INFORMATION:
; APPLICANT: Jerome B. Posner
; APPLICANT: Josep O. Dalmau
; APPLICANT: Myrna R. Rosenfeld
; TITLE OF INVENTION: Ma FAMILY POLYPEPTIDES AND ANTI-Ma
; TITLE OF INVENTION: ANTIBODIES
; FILE REFERENCE: 2581.1004-004
; CURRENT APPLICATION NUMBER: US/10/037,860
; CURRENT FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 09/189,527
; PRIOR FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 329
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-037-860-4

Query Match      31.6%; Score 766.5; DB 13; Length 329;
Best Local Similarity 50.2%; Pred. No. 1.6e-63;
Matches 157; Conservative 50; Mismatches 103; Indels 3; Gaps 2;

Qy      1  MPTLLQDWCRCGEHLNTRRCMLILGIPEDCGDEFEETLQACRHLGRYRVIQGRMFRREE 60
Db      1  MAMTLLEDWCRCGMDVNSQRTLLVWGIPVNCDEAEIETLQAAMPQVS-YRMLGRMFWREE 59

Qy      61  NQAAILLELAQDIDYALLPREIPKGGPWEVIVKPRNSDGEF--LNRLNRFLEERRTVS 120
Db      60  NAKAALLELTGAVDYAAIPREMPGKGGVWKVLFKPPSTDAAEFLERLHLFLAREGWTVDV 119

Qy      121  NRVLGSDTNCSPRVTTISPEFTWTAQTLGAAVQPLLEQMLYRELRFVSGNTISIPGALAF 180
Db      120  ARVLGFQNFPTTPGPEMPAEMLYN--ILDNVIQPLVESIWKRLTILFSGKGHPRAWGNF 177

Qy      181  DAMLEHTTEMLQWQVPEGEKRRRLMECLRGPAQVQVSGLRASNASITVBECLAALQOVF 240
Db      178  DPWLEHTNEVLEWQVSDVEKRRLEMSLRGPAADVIRILKSNPNPAITTAELKALEQVF 237

Qy      241  GPVESHKIAQVKLCAYQAEAGKVSFVLRLEPLLQRAVENNVVSRNNVQTRLKRVLSG 300
Db      238  GSVESSRDAQIKFLNTYCNFEKLSAYVIRLEPLLQKVVEKGAIDKDNVNQARLEQVIAG 297

Qy      301  ATLPDKLRDKLKL 313
Db      298  ANHSGAIRRQLWL 310

RESULT 11
US-09-804-014A-40
; Sequence 40, Application US/09804014A
; Publication No. US20030064489A1
; GENERAL INFORMATION:
; APPLICANT: Li, Li
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Vernet, Corine
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard
; APPLICANT: Spaderna, Steven
; APPLICANT: Majumder, Kumud
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-721 US
; CURRENT APPLICATION NUMBER: US/09/804,014A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/188,316
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/188,277
; PRIOR FILING DATE: 2000-03-10
```



```
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-804-014A-73

Query Match      30.6%; Score 740.5; DB 12; Length 312;
Best Local Similarity 48.8%; Pred. No. 4.2e-61;
Matches 148; Conservative 51; Mismatches 101; Indels 3; Gaps 2;

Qy 1 MPTLLQDWCRGHEHLNTRRCMLILGIPEDCGEDEFETLQEAACHLGRYVIGRMFRREE 60
Db 1 MTLRLLEDWCRGMDMPKALLIAGISQSCVAIEIEALQAGLAPLGEYLLGRMFRDE 60

Qy 61 NQAAILLELAQDIDYALLPREIPKGGPWEVIVKPRNSDGEFLNRLNRFLEERRTVSDM 120
Db 61 NRKVALVGLTAETSHALVPKEIPKGGIWRVIFKPPDPDNTFLSRLNEFLAGEGTMVTEL 120

Qy 121 NRVLGSDTNCSPARVTISPEFWT--WAOTLGAAVQPLLEOMLYRELRFVSGNTISIPGAL 178
Db 121 SRALGHENGSLDPEQGMIPEMWAPMLAQAL--EALQPALQCLKYKLVFSGRESPEEGEE 179

Qy 179 AFDWLHTTEMQLQMWQVPEGEKRRRLMECLRGPAQVSGLRASNASITVEECIAAQQ 238
Db 180 EFRGWFHTTQMIKAWQVDPVEKRRRLLESIRGPAQVIRVILKINPLIITVDECLQALEE 239

Qy 239 VFGPVESHKIAQVKLCKAYQEAQKVSFVLRLEPLLQRAVENNVSRNNVQTRLKRVL 298
Db 240 VFGVTDNPRELQVKYLTYYQKDEKLSAYVLRLEPLLQKLVQGAIERDAVNQARLDQVI 299

Qy 299 SGA 301
Db 300 AGA 302

RESULT 14
US-09-804-014A-74
; Sequence 74, Application US/09804014A
; Publication No. US20030064489A1
; GENERAL INFORMATION:
; APPLICANT: Li, Li
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Vernet, Corine
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard
; APPLICANT: Spaderna, Steven
; APPLICANT: Majumder, Kumud
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-721 US
; CURRENT APPLICATION NUMBER: US/09/804,014A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/188,316
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/188,277
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/189,139
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/189,140
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/190,401
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/190,231
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 74
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-804-014A-74

Query Match      30.6%; Score 740.5; DB 12; Length 312;
Best Local Similarity 48.8%; Pred. No. 4.2e-61;
Matches 148; Conservative 51; Mismatches 101; Indels 3; Gaps 2;
```

```
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-037-860-11

Query Match      25.5%; Score 618.5; DB 13; Length 283;
Best Local Similarity 50.2%; Pred. No. 1.2e-49;
Matches 135; Conservative 44; Mismatches 85; Indels 5; Gaps 3;

Qy 82 IPKGGPWEVIVKPRNSDGEFLNRLNRFLEERRTVSDMNRVLGSDTNCSPARVTISPEF 141
Db 1 VOGKGGVWKVIFKTPNQDTEFLERLNLFLKEGQTVSGMFRALGQGVSPATVPCISPEL 60

Qy 142 WT--WAOTLGAAVQPLLEOMLYRELRFVSGNTISIPGALAFDAMLEHTTEMLQWQVPEG 199
Db 61 LAHLGQAMAHAPQPLL--PMRYKLRVFGSAVPAPEESFEVWLEQATEIVKEMPVTEA 119

Qy 200 EKRRRLMECLRGPAQVSGLRASNASITVEECIAAQQVQGVESHKIAQVKLCKAYQE 259
Db 120 EKKWLAESLGRPALDLMHIVQADNPISLSVEECLEAFQVFGSLESRRTAQVRLKYQE 179

Qy 260 AGEKVSFVLRLPLQRAVENNVSRNNVQTRLKRVLSGATLPDKLRDKLKLKMKQRK 319
Db 180 EGEKVSAYVLRLETLRKXAKRAIPRIADQVRLQWAGATLNQMLWCLRRLKQDGP 239

Qy 320 PPGFIALVKLRREBEWEATLGPDRSLE 348
Db 240 PPSFLELMKVIREBEWEASF--ENESIE 266

RESULT 15
US-10-037-860-11
; Sequence 11, Application US/10037860
; Publication No. US20020123114A1
; GENERAL INFORMATION:
; APPLICANT: Jerome B. Posner
; APPLICANT: Joseph O. Dalmat
; APPLICANT: Myrna R. Rosenfeld
; TITLE OF INVENTION: Ma FAMILY POLYPEPTIDES AND ANTI-Ma
; TITLE OF INVENTION: ANTIBODIES
; FILE REFERENCE: 2581.1004-004
; CURRENT APPLICATION NUMBER: US/10/037,860
; CURRENT FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 09/189,527
; PRIOR FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 283
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-037-860-11

Query Match      25.5%; Score 618.5; DB 13; Length 283;
Best Local Similarity 50.2%; Pred. No. 1.2e-49;
Matches 135; Conservative 44; Mismatches 85; Indels 5; Gaps 3;
```

Search completed: September 27, 2004, 17:26:07
Job time : 137.663 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 27, 2004, 17:06:36 ; Search time 23.4526 Seconds
(without alignments)
1019.197 Million cell updates/sec

Title: US-10-037-860-13
Perfect score: 2423
Sequence: 1 MPTLLQDWCGRGHLNTRRC.....VESGNGNWMDKSHPKSKAK 463

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A-COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B-COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A-COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B-COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS-COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2394	98.8	462	4	US-09-189-527-13
2	766.5	31.6	329	4	US-09-189-527-4
3	462.5	19.1	195	4	US-09-189-527-7
4	122	5.0	2293	3	US-09-368-530-2
5	110.5	4.6	341	4	US-09-252-991A-20182
6	110	4.5	718	4	US-09-252-991A-32743
7	104	4.3	383	4	US-09-489-039A-11848
8	101.5	4.2	1201	4	US-09-252-991A-32259
9	101	4.2	580	4	US-09-252-991A-22036
10	100	4.1	363	4	US-09-252-991A-26726
11	100	4.1	369	4	US-09-252-991A-22549
12	99.5	4.1	1050	4	US-09-555-554-2
13	99	4.1	639	1	US-08-466-390-2
14	99	4.1	639	1	US-08-470-950-2
15	99	4.1	639	1	US-08-467-781-2
16	99	4.1	639	1	US-08-195-487-2
17	99	4.1	639	2	US-08-483-924-2
18	99	4.1	639	5	PCT-US93-06160-2
19	98.5	4.1	733	4	US-09-489-039A-12568
20	98.5	4.1	820	4	US-09-252-991A-23346
21	97.5	4.0	776	4	US-09-252-991A-28446
22	97.5	4.0	2154	2	US-08-841-349-4
23	97.5	4.0	2154	4	US-09-431-184A-4
24	97	4.0	181	4	US-09-252-991A-26482
25	97	4.0	541	4	US-09-252-991A-30867
26	97	4.0	1958	1	US-07-945-283-2
27	97	4.0	2108	4	US-09-252-991A-31502

28	96.5	4.0	406	4	US-09-252-991A-24973	Sequence 24973, A
29	96.5	4.0	771	4	US-09-252-991A-20455	Sequence 20455, A
30	96.5	4.0	773	4	US-09-913-301-5	Sequence 5, Appli
31	96.5	4.0	804	4	US-09-913-301-2	Sequence 2, Appli
32	96.5	4.0	849	4	US-09-252-991A-17953	Sequence 17953, A
33	96.5	4.0	920	4	US-09-252-991A-28918	Sequence 28918, A
34	96	4.0	997	4	US-09-252-991A-30799	Sequence 30799, A
35	95.5	3.9	371	4	US-09-252-991A-22517	Sequence 22517, A
36	95.5	3.9	577	4	US-09-489-039A-9575	Sequence 9575, Ap
37	95.5	3.9	668	1	US-08-205-018-2	Sequence 2, Appli
38	95.5	3.9	859	1	US-08-395-580-2	Sequence 2, Appli
39	95.5	3.9	859	5	PCT-US95-02792-2	Sequence 2, Appli
40	95.5	3.9	1041	4	US-09-252-991A-25637	Sequence 25637, A
41	95.5	3.9	1053	4	US-09-252-991A-26140	Sequence 26140, A
42	95	3.9	390	4	US-09-252-991A-22732	Sequence 22732, A
43	95	3.9	634	1	US-08-339-152A-17	Sequence 17, Appl
44	95	3.9	653	1	US-08-339-152A-16	Sequence 16, Appl
45	95	3.9	653	2	US-08-007-999B-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
; Sequence 13, Application US/09189527A
; Patent No. 6387639
; GENERAL INFORMATION:
; APPLICANT: Jerome B. Posner
; APPLICANT: Josep O. Dalmau
; APPLICANT: Myrta R. Rosenfeld
; TITLE OF INVENTION: Ma Family Polypeptides and Anti-Ma
; FILE REFERENCE: SLK98-01
; CURRENT APPLICATION NUMBER: US/09/189,527A
; CURRENT FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 462
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-189-527-13

Query Match	98.8%	Score	2394	DB	4	Length	462
Best Local Similarity	100.0%	Pred. No.	3e-247				
Matches	457	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
Qy	7	QDWCGRGHLNTRRCMLILGIPEDCGEDEFBETLOEACRHILGRYVIGRMFRERENQAAIL	66				
Db	1	QDWCGRGHLNTRRCMLILGIPEDCGEDEFBETLOEACRHILGRYVIGRMFRERENQAAIL	60				
Qy	67	LELAQDIDYALLPREIPGKGPMEVIYKPRNSDGEFLNRLNRFLEEBERRTVSDMNRVLGS	126				
Db	61	LELAQDIDYALLPREIPGKGPMEVIYKPRNSDGEFLNRLNRFLEEBERRTVSDMNRVLGS	120				
Qy	127	DTNCSAPRVITSPBFTWAOITLGAAVQPLLEQMLYRELRFVPSGNTISIPGALAFDAMLEH	186				
Db	121	DTNCSAPRVITSPBFTWAOITLGAAVQPLLEQMLYRELRFVPSGNTISIPGALAFDAMLEH	180				
Qy	187	TTEMLOWQVPEGKRRRLMECLRGPAOVVSGLRASNASITVEECALAOQVFGPVESH	246				
Db	181	TTEMLOWQVPEGKRRRLMECLRGPAOVVSGLRASNASITVEECALAOQVFGPVESH	240				
Qy	247	KIAQVKCKAYQAGEKVSFVLRLPLELQRAVENNVSRNNVQTRLKRVLGSATLPDK	306				
Db	241	KIAQVKCKAYQAGEKVSFVLRLPLELQRAVENNVSRNNVQTRLKRVLGSATLPDK	300				
Qy	307	LRDKLKLKMKQRKPPGFLALVKLLREEEWEATLGPDRSLEGLVAPRPPARITGVAV	366				
Db	301	LRDKLKLKMKQRKPPGFLALVKLLREEEWEATLGPDRSLEGLVAPRPPARITGVAV	360				
Qy	367	PLFASGNSFARPSQGYRRRRGRGQHRRGGVARAGSRGKRKHHTFCYSCGEDGHRVQ	426				


```
Db 361 PIPASGNSFARPSQYRRRGGRGQHRGVARAGSRGKRKRHTFCISCGEDGHIRVQ 420
QY 427 CINPSNLLVKKQKAAVESGNGNWNWADKSHPKSKAK 463
Db 421 CINPSNLLVKKQKAAVESGNGNWNWADKSHPKSKAK 457
```

RESULT 2

```
US-09-189-527-4
; Sequence 4, Application US/09189527A
; Patent No. 6387639
; GENERAL INFORMATION:
; APPLICANT: Jerome B. Posner
; APPLICANT: Josep O. Dalmau
; APPLICANT: Myrna R. Rosenfeld
; TITLE OF INVENTION: Ma Family Polypeptides and Anti-Ma
; TITLE OF INVENTION: Antibodies
; FILE REFERENCE: SLK98-01
; CURRENT APPLICATION NUMBER: US/09/189,527A
; CURRENT FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 329
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-189-527-4
```

```
Query Match 31.6%; Score 766.5; DB 4; Length 329;
Best Local Similarity 50.2%; Pred. No. 2.9e-73;
Matches 157; Conservative 50; Mismatches 103; Indels 3; Gaps 2;

QY 1 MPTLLQDWCRCGHLNTRCMLILGIPEDCGEDEFETLQACRHLGRYRVIGRMFRREE 60
Db 1 MAMTLLDWCRCGMDVNSQRTLLVWGPVNCDEAIEETLQAAMPQVS-YRMLGRMFWRRE 59

QY 61 NQAAILLEAQDIDYALLPREIPKGGPWEVIVKPRNSDGEFLNRLNRELEERTVSOM 120
Db 60 NAKAALLEITGAVDYAAIPREMPKGGVWKVLPKPTSDAEFLRLHLFLAREGWVQDV 119

QY 121 NRVLGSDTNCSPRVVTISPEFTWQAOTLGAAGVPLLEQMLYRLRVFSGNTISIPGALAF 180
Db 120 ARVLGQNTPTPGPMPAEMLN--LLDNVIQPLVESIWKLTIFSGKGHPWRANGF 177

QY 181 DAWLEHTTEMLQWQVPEGEKRRRLMECLRGPALQVWSGLRASNASITVEECLAAALQQVF 240
Db 178 DPWLEHTNEVLBEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEQV 237

QY 241 GPVESHKIAQVKLCKAYQAGEKVSFVLRLLEPLLQRAVENVVSRNNVOTRLKRVLSG 300
Db 238 GSVESSRDAQIKFLNTYQNGEKLISAYVIRLEPLLQKRVKGAIDKONVNOARLEQVIAG 297

QY 301 ATLPDKLRDKLXL 313
Db 298 ANHSGAIRQLWL 310
```

RESULT 3

```
US-09-189-527-7
; Sequence 7, Application US/09189527A
; Patent No. 6387639
; GENERAL INFORMATION:
; APPLICANT: Jerome B. Posner
; APPLICANT: Josep O. Dalmau
; APPLICANT: Myrna R. Rosenfeld
; TITLE OF INVENTION: Ma Family Polypeptides and Anti-Ma
; TITLE OF INVENTION: Antibodies
; FILE REFERENCE: SLK98-01
; CURRENT APPLICATION NUMBER: US/09/189,527A
; CURRENT FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
```

```
; SEQ ID NO 7
; LENGTH: 195
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-189-527-7

Query Match 19.1%; Score 462.5; DB 4; Length 195;
Best Local Similarity 47.7%; Pred. No. 4.2e-41;
Matches 93; Conservative 34; Mismatches 65; Indels 3; Gaps 2;

QY 2 PTLTLDWCRCGHLNTRCMLILGIPEDCGEDEFETLQACRHLGRYRVIGRMFRREEN 61
Db 1 PLALLDWCRCIMSDQKSLMVTGIPADPEEAIEQVLTLLKSLGRYLLGKIFRKQEN 60

QY 62 AQAAILLEAQDIDYALLPREIPKGGPWEVIVKPRNSDGEFLNRLNRELEERTVSOM 121
Db 61 ANAVILLELEDDTDSAIPSEVQKGGVWKVIFKTPNQDTEFLERLNLFLKEKGQTVSGMF 120

QY 122 RVLGSDTNCSPRVVTISPEFTW--WAOTLGAAGVPLLEQMLYRLRVFSGNTISIPGALA 179
Db 121 RALGQEAALSPATVPCISPELLAHLGQAMAHAPQPLL-PMRYKRLRVFSGSAVPAPFEES 179

QY 180 FDWLEHTTEMLQW 194
Db 180 FEVWLEQATEIVKEW 194
```

RESULT 4

```
US-09-368-590-2
; Sequence 2, Application US/09368590
; Patent No. 6187563
; GENERAL INFORMATION:
; APPLICANT: Solimena, Michele
; TITLE OF INVENTION: INTERACTING POLYPEPTIDES FOR
; FILE REFERENCE: 101918-200 (OCR-941)
; CURRENT APPLICATION NUMBER: US/09/368,590
; CURRENT FILING DATE: 1999-08-04
; EARLIER APPLICATION NUMBER: 60/095,657
; EARLIER FILING DATE: 1998-08-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 2293
; TYPE: PRT
; ORGANISM: Human
US-09-368-590-2
```

```
Query Match 5.0%; Score 122; DB 3; Length 2293;
Best Local Similarity 25.1%; Pred. No. 0.0069;
Matches 100; Conservative 57; Mismatches 144; Indels 98; Gaps 24;

QY 60 ENAQAILLEAQDIDYALLPREIPK-----GGFWEVIVKPRNSDGEFL---NR 105
Db 733 EPRQAALLEEA-----ALLAERFPAQAAXLHQGAEBLGAEWGALASAAACGGAFAAAGR 787

QY 106 LNRFLEERTVSDMNRVLGSDTNCSPRVVTISPEFTWQAOTLGAAGVPLLEQMLYRE-- 163
Db 788 LQRFLEHDLDAFLDQWVRAQEAAGSGSGP---LPNSLEADALLAHLAKKEEDOREED 843

QY 164 -LRVFSGN--TISPGA-----LAFDAWLEHTT-----EMLQWQVPEGEKRRRLMEC--- 208
Db 844 YARIVAASEALLAADAAGAEELGPGALDEWLPHELGWHKLLGLMKA---RRKALVQAHIY 899

QY 209 ---LRG--PALQVWSG--LRASNASI--TVEECLAAALQVFGPVESHKIAQVKLCKAYQE 259
Db 900 QLFRLDRQALVVLNRQEMALSGAELPGTVESVEEALKQHRDPLTTMELSSQQQKQVAVQA 959

QY 260 A-----GEKVSSFVLRLEPLLQRAVENVVSRNV---NOTRLKRVLS-----G 300
Db 960 AEGLLRQGNITYEQAEQAVTR---LLEKQENQLRAQQWQKLDHQLQHLRDCHDCHELD 1016

QY 301 ATLPDKL-----BDKLLMKQRKPPGFLALVKLLREERWEATL-----GPDRESLE- 348
```

```
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20182
Query Match      4.6%; Score 110.5; DB 4; Length 341;
Best Local Similarity 24.3%; Pred. No. 0.005;
Matches 66; Conservative 22; Mismatches 109; Indels 75; Gaps 11;

Db 1017 GWIHEKMLMARDGTREDNHKLKWLKHQAQFMA---ELAQNKWEKIEREGPATDAGEA 1073
Qy 349 ---GL--EVAPRPPARITGVGA-VPLPASGNSFDARPSQ 381
Db 1074 RTGGLRAEEAGRDPPVLGGAGEHHPGRHGSSL--RPSK 1110

RESULT 5
US-09-252-991A-20182
; Sequence 20182, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20182
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20182

Query Match      4.6%; Score 110.5; DB 4; Length 341;
Best Local Similarity 24.3%; Pred. No. 0.005;
Matches 66; Conservative 22; Mismatches 109; Indels 75; Gaps 11;

Qy 208 CLRGPALQVSLGRASNASITVECLAAALQOVFGFVESHKIAQVKLCKA-----256
Db 16 CLGCP-----VDGHPAKHASLRLGNGNKLHGVPAQH--RLCFAGGGQPRDPAH 67
Qy 257 -----YQEA-----GEKVSSFVLRLEPLLRQAVENNVVSRNNVQTRLKRVLSGATLPDX 306
Db 68 PPAQPHYSSHPRHGRPALPALRADPLRLRSGR--SORRCPCTRQ---AEAADFDR 122
Qy 307 LRDLKLMKORRKPGLFALVKLREBEWEATLGPRESLE-----348
Db 123 HRPALRDPHRLRPP---AVPRDLRDPHGQPRAGPARGLRRGHRGLRAARFELHPPA 179
Qy 349 -----GLEVAPRPPARITGVGAVPLPASGNSFDARPSQGYRR--RRGRGQHRR 394
Db 180 HRRDLQHSRLAGVPAPERHAERAFAGAGQPRIPAPGQPGGAAGSLAVRSORPGNGLHRR 239
Qy 395 -----GGVARAGSRGRKRTFCVSCGEDGH 422
Db 240 LAVPGQCRRRDRGNPLRTGHR--HPAGVFGH 269

RESULT 6
US-09-252-991A-32743
; Sequence 32743, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32743
; LENGTH: 718
```

```
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32743
Query Match      4.5%; Score 110; DB 4; Length 718;
Best Local Similarity 21.7%; Pred. No. 0.02;
Matches 70; Conservative 31; Mismatches 93; Indels 128; Gaps 16;

Qy 242 PVESHKIAQVKLCKAYQAGEKVSFVLRLEPL-----LQRAVE-----280
Db 122 PEKPHHVERPRHVDAYRPGQQRATR---RLSPFGPRRRRTDRHAGGGPQRAHLHLARA 178
Qy 281 -----NNVSGRRNVNQTRLKRVLSGATLPDKLRLKLMKQRRKPPG---FLALV 327
Db 179 DGHRRCLHPLRPAGRRGLPRRLRGAV--ANLPSRADH---GVRRRPAGGLFRGLA 232
Qy 328 KLLREBEWEATLGPRESLEGLEVAAPRPPAR-ITG-----VGAVPLPASGNSF 375
Db 233 QRTVPGDAARTRPAADRDRRLRRRAPARGRTGRGGGGRQPLPAGAIQPA-----287
Qy 376 DARPSQGYRRRG-----RQHR-----393
Db 288 -GEPPQGLQRRPCPPAFPPRSAGGARGHLRGLHRRPAPVRRRARPGRSQKRVAAG 346
Qy 394 RGVVARAGSRGRK-----RKRTFCVSCGEDGHIRVQCINPSNLLL-----VKQK 439
Db 347 RGGPRAGLPAHRRRHGFLPRRHSYRPA---DGH-RAGLPGPRSOYLOGAPGDPVHR 402
Qy 440 KQAAVESGNGNMAWDKSHPKSK 461
Db 403 AHQPAPGGRG-----HPRPR 417

RESULT 7
US-09-489-039A-11848
; Sequence 11848, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11848
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11848

Query Match      4.3%; Score 104; DB 4; Length 383;
Best Local Similarity 22.5%; Pred. No. 0.03;
Matches 60; Conservative 44; Mismatches 91; Indels 72; Gaps 12;

Qy 114 RRTVSDMNRVLGSDTNCASAPRTVISPEFTWAQTLGAAVQPLLEQMLYRELURVFSGNTIS 173
Db 118 QRGAAELHR-CRSETLCTLRATISSSFITHPRNNNDIAHNLAQ---VRD-----K 164
Qy 174 IPGALAPDAWLEHTTEMLQMWQVPEGEKRRRLMECLRGPALQVVSGLRASNASITVECL 233
Db 165 ISGAAA-----RCGRAPEEVTLLAVSKTKPASAIEAI 197
Qy 234 AALQOVFGFVESHKIAQVKLCKAYQAGEKVSFVLRLEPLQRA-----VENNVVSRNV 289
Db 198 AAGQRAFQ--ENV-----VQEGVEKINH-----QAGVSLGWLHFGIPLOS 237
Qy 290 NOTRL-KRVLSGATLPDKLRLKLMKQRRKPPGFLALVKLLRE--EEWEATLGPRES 346
Db 238 NKSRLVAEHFDWCHTVDRLKTIATRLNEQR---PAHLPLKVLIIQINISDEQSKSIGIPLEA 294
```



```

; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-467-781-2

Query Match      4.1%; Score 99; DB 1; Length 639;
Best Local Similarity 20.8%; Pred. No. 0.24;
Matches 79; Conservative 57; Mismatches 162; Indels 82; Gaps 17;

Qy 68 ELAQDIDYALLPREIPGKGPEWVIVKPRNSDGEFLNRLNRFLEEEERTVSDMNRVLGSD 127
Db 250 DEKRELD-SITPEVLPG-----WKGMSVSDLADKLSTDLSLJAHARRIDQLNRELAEQ 304
Qy 128 TNCAPRVITISPEFTWAQ--TLGAAVQPLLE-----QMLYRELRFVSGNTISIP 175
Db 305 KATEKQHITLAEKQKLEEKRAFDSDAVAKALEHHRSEIQAEQDRKIEBEVRDAMENEMRTP 364
Qy 176 GALAFDAWLEHTTMTLQMWQVPEGEKR---RRLMECLRGFPALQV--VSGLRASNASITV 229
Db 365 SPTA-----AAHTDHLRDVLRVQEQELKSEFEQNLSEKLSQEQLQFRRLSQEQVDNFTLDI 420
Qy 230 EECALAQVFGPVESHKIAQVLCCKAYQ-----EAGEKVSSFVLRLEPLLQRAVEN 281
Db 421 NTAYARLRGIEQAVQSHAVAEAEARKAHQLWLSVEALKYSMTSSAETPTIP-LGSAVE- 478
Qy 282 NVYSRRNVNQLKRVLSGATLPDKLRDKL---KLMQQRKPPGFLALVKLLREEEWEA 338
Db 479 --AIKANCSDNEFTQALTAIIPPESLTRGVYSEETLRAR-----FYAVQKLARR-----VA 527
Qy 339 TLGPDRES-----LEGLEVAP---RPPARITGVGAVPLPASGNSFDARPSQGYRRR 386
Db 528 MIDETRNSLYQYFLSYLSQLLLFPFPQQLKPPPEL-----CPEDINTFKLLSYASYCIE 580
Qy 387 RG-----RGQHR 394
Db 581 HGDLELAAKFVNQLKGSRR 600

```

Search completed: September 27, 2004, 17:11:20
Job time : 25.4526 secs

Blank

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 27, 2004, 17:02:20 ; Search time 77.9232 Seconds
(without alignments)
1678.826 Million cell updates/sec

Title: US-10-037-860-13
Perfect score: 2423
Sequence: 1 MPTLLQDWCRGEHLNTRRC.....VESGNGWAWDKSHPKSKAK 463

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_29Jan04:*

- 1: geneseqp1980s:*
- 2: geneseqp1980s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2423	100.0	463	3 AAB12529	Human Ma5
2	2423	100.0	463	3 AAB42315	Human ORF
3	2283.5	94.2	455	5 ABB05727	Human sig
4	883.5	36.5	452	6 AAO16179	Human pro
5	836.5	34.5	364	7 ADC08977	Onconeuro
6	818	33.8	353	3 AAB74701	Human mem
7	768.5	31.7	351	4 AAE01340	Human gen
8	768.5	31.7	351	4 AAB74695	Human mem
9	768.5	31.7	351	4 AAU08664	Human NOV
10	768.5	31.7	399	6 ADA54410	Human pro
11	768.5	31.7	399	6 ABG99947	Human nov
12	755.5	31.2	329	3 AAB12525	Human ORF
13	740.5	30.6	312	3 AAB43023	Human ORF
14	617.5	25.5	283	4 AAB12528	Human Ma4
15	476.5	19.7	280	4 AAE01336	Human gen
16	462.5	19.1	195	3 AAB12526	Human Ma2
17	454.5	18.8	237	4 AAB94854	Human pro
18	429.5	17.7	403	5 AAM51624	KIAA0883-
19	423.5	17.5	403	5 ABG97495	Human NOV
20	423.5	17.5	403	6 ABO14772	Human hum
21	421.5	17.4	403	6 ABO14773	Novel hum
22	415.5	17.1	402	4 AAB60478	Human cel
23	384.5	15.9	337	4 AAM25693	Human pro
24	359	14.8	80	4 ABB15036	Human ner
25	351.5	14.5	439	6 ABP75736	Human sec

26	323	13.3	149	3 AAB12527	Human Ma3
27	270	11.1	116	4 AAM14937	Peptide #
28	270	11.1	116	4 ABB33906	Peptide #
29	270	11.1	116	4 AAM27367	Peptide #
30	270	11.1	116	4 ABB28727	Peptide #
31	270	11.1	116	4 ABB19347	Protein #
32	270	11.1	116	4 AAM67075	Human Bon
33	270	11.1	116	4 AAM54673	Human bra
34	270	11.1	116	4 ABG48741	Human liv
35	270	11.1	116	4 AAM02665	Peptide #
36	270	11.1	116	5 ABG36734	Human pep
37	223.5	9.2	615	4 ABG19651	Novel hum
38	223.5	9.2	615	4 ABG14259	Novel hum
39	218	9.0	110	4 AAO01787	Human pol
40	191	7.9	538	4 ABUS2641	Human bra
41	191	7.9	538	7 ADC31124	Human nov
42	155	6.4	584	4 AAU28186	Novel hum
43	124	5.1	557	7 ADC72914	HIV-1 gag
44	122	5.0	2293	3 AAY69197	Amino aci
45	113.5	4.7	473	4 ABG12120	Novel hum

ALIGNMENTS

RESULT 1

AAB12529
ID AAB12529 standard; protein; 463 AA.

XX AC AAB12529;

XX DT 02-NOV-2000 (first entry)

XX DE Human Ma5 protein SEQ ID NO:13.

XX KW Ma1; Ma2; Ma3; Ma4; Ma5; paraneoplastic limbic encephalitis; diagnosis;
paraneoplastic syndrome; neoplasm; colon cancer; brainstem encephalitis;
breast cancer; parotid gland cancer; lung cancer; testicular cancer;
germ-cell tumour.

XX OS Homo sapiens.

XX PN JP2000146982-A.

XX PD 26-MAY-2000.

XX PF 10-NOV-1999; 99JP-00320171.

XX PR 10-NOV-1998; 98US-00189527.

XX PA (SLOK) SLOAN KETTERING INST CANCER RES.

XX WPI; 2000-468119/41.

XX N-PSDB; AAA60837.

PT Diagnosing paraneoplastic syndrome e.g. paraneoplastic limbic
encephalitis or neoplasm e.g. colon cancer comprising assessing a test
sample for the presence or absence of antibodies to a Ma family
polypeptide.

XX Claim 48; Fig 9-10; 27pp; Japanese.

CC The present invention describes a method for diagnosing a paraneoplastic
syndrome or neoplasm. The method comprises assessing a test sample for
the presence or absence of antibodies to a Ma family polypeptide (I). The
method is used to diagnose a paraneoplastic syndrome especially
paraneoplastic limbic encephalitis and/or brainstem encephalitis or
neoplasm in an individual. The method diagnoses the neoplasm by assessing
antibodies to (i) preferably Ma1, which is indicative presence of breast
cancer, colon cancer, parotid gland cancer, lung cancer, testicular
cancer, and germ-cell tumours or Ma2, which is indicative of testicular
cancer, germ-cell tumour, and lung cancer. The present sequence is the
Ma5 protein as given in the present invention

```
XX SQ Sequence 463 AA;
Query Match 100.0%; Score 2423; DB 3; Length 463;
Best Local Similarity 100.0%; Pred. No. 5.6e-231;
Matches 463; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPTLLQDWCRGHEHLNTRCMLILGIPEDCGEDEFETLQEACRHLYGRYRVIGRMFRREE 60
DB 1 MPTLLQDWCRGHEHLNTRCMLILGIPEDCGEDEFETLQEACRHLYGRYRVIGRMFRREE 60
QY 61 NQAAILLELAQDIDYALLPREIPGKGPWEIVKPRNSDGEFLNRLNRFLEBERRTVSDM 120
DB 61 NQAAILLELAQDIDYALLPREIPGKGPWEIVKPRNSDGEFLNRLNRFLEBERRTVSDM 120
QY 121 NRVLGSDTNCSPRVVTISPEFTWTAQTGAAVQPLLEQMLYRELRFVSGNTTISIPGALAF 180
DB 121 NRVLGSDTNCSPRVVTISPEFTWTAQTGAAVQPLLEQMLYRELRFVSGNTTISIPGALAF 180
QY 181 DAWLEHTTLMQWQVPEGEKRRRLMECLRGPAQVVSGLRASNASITVEECALAQOVF 240
DB 181 DAWLEHTTLMQWQVPEGEKRRRLMECLRGPAQVVSGLRASNASITVEECALAQOVF 240
QY 241 GPVESHKIAQVKLCKAYQAGEKVSFVLRLPELLOQAVENNVSRNNVQTLKRVLSG 300
DB 241 GPVESHKIAQVKLCKAYQAGEKVSFVLRLPELLOQAVENNVSRNNVQTLKRVLSG 300
QY 301 ATLPDKLRDKLKMQRKPPGFLALVKLLREBEEWEATLGPDRSLEGLVAPRPARI 360
DB 301 ATLPDKLRDKLKMQRKPPGFLALVKLLREBEEWEATLGPDRSLEGLVAPRPARI 360
QY 361 TGVGAVPLPASGNSFDARPSQGVRRRGRGOHRGGVARAGSGRKRKHTFCYSCGD 420
DB 361 TGVGAVPLPASGNSFDARPSQGVRRRGRGOHRGGVARAGSGRKRKHTFCYSCGD 420
QY 421 GHIRVQCINPSNLLLVKQKQAAVESGNGNWADKSHPKSKAK 463
DB 421 GHIRVQCINPSNLLLVKQKQAAVESGNGNWADKSHPKSKAK 463
RESULT 2
AAB42315
ID AAB42315 standard; protein; 463 AA.
XX AC AAB42315;
XX DT 08-FEB-2001 (first entry)
XX DE Human ORFX ORF2079 polypeptide sequence SEQ ID NO:4158.
XX KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
XX KW vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
XX KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
XX KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
XX KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
XX KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
XX KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
XX KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
XX KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
XX KW cholesterol ester storage; systemic lupus erythematosus; infection;
XX KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
XX KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
XX KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
XX KW thrombosis; contraceptive.
XX OS Homo sapiens.
XX PN WO200058473-A2.
XX XX
XX XX
XX XX
XX XX
XX PF 31-MAR-2000; 2000WO-US0008621.
XX XX
```

```
PR 31-MAR-1999; 99US-0127607P.
PR 02-APR-1999; 99US-0127636P.
PR 05-APR-1999; 99US-0127728P.
PR 30-MAR-2000; 2000US-00540763.
XX (CURA-) CURAGEN CORP.
XX PA
XX PI Shinkets RA, Leach M;
XX XX
XX WPI; 2000-602362/57.
XX DR N-PSDB; AAC76524.
XX XX
XX Novel nucleic acids and peptides derived from open reading frame X,
XX useful for treating e.g cancers, proliferative disorders,
XX neurodegenerative disorders and cardiovascular disease.
XX
XX Claim 11; Page 3345-3347; 5507pp; English.
XX
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
XX which represent the human ORFX open reading frames 1 to 3161. The ORFX
XX sequences have activities such as: cytostatic; hepatotropic; vulnery;
XX antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
XX anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
XX cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
XX dermatological; immunosuppressive; antiinflammatory; antibacterial;
XX antiviral; antifungal; antirheumatic; antithyroid; and antianaemic. The
XX sequences can be used for determining the presence of or predisposition
XX to, or preventing or treating pathological conditions associated with an
XX ORFX-associated disorder. The nucleic acids can be used to express ORFX
XX proteins in gene therapy vectors. The proteins and nucleic acids may be
XX used to treat cancers, proliferative disorders, neurodegenerative
XX disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
XX diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
XX storage, systemic lupus erythematosus, severe combined immunodeficiency
XX (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
XX disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
XX cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
XX enhance coagulation; to inhibit thrombosis; and as a contraceptive
XX
XX SQ Sequence 463 AA;
Query Match 100.0%; Score 2423; DB 3; Length 463;
Best Local Similarity 100.0%; Pred. No. 5.6e-231;
Matches 463; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPTLLQDWCRGHEHLNTRCMLILGIPEDCGEDEFETLQEACRHLYGRYRVIGRMFRREE 60
DB 1 MPTLLQDWCRGHEHLNTRCMLILGIPEDCGEDEFETLQEACRHLYGRYRVIGRMFRREE 60
QY 61 NQAAILLELAQDIDYALLPREIPGKGPWEIVKPRNSDGEFLNRLNRFLEBERRTVSDM 120
DB 61 NQAAILLELAQDIDYALLPREIPGKGPWEIVKPRNSDGEFLNRLNRFLEBERRTVSDM 120
QY 121 NRVLGSDTNCSPRVVTISPEFTWTAQTGAAVQPLLEQMLYRELRFVSGNTTISIPGALAF 180
DB 121 NRVLGSDTNCSPRVVTISPEFTWTAQTGAAVQPLLEQMLYRELRFVSGNTTISIPGALAF 180
QY 181 DAWLEHTTLMQWQVPEGEKRRRLMECLRGPAQVVSGLRASNASITVEECALAQOVF 240
DB 181 DAWLEHTTLMQWQVPEGEKRRRLMECLRGPAQVVSGLRASNASITVEECALAQOVF 240
QY 241 GPVESHKIAQVKLCKAYQAGEKVSFVLRLPELLOQAVENNVSRNNVQTLKRVLSG 300
DB 241 GPVESHKIAQVKLCKAYQAGEKVSFVLRLPELLOQAVENNVSRNNVQTLKRVLSG 300
QY 301 ATLPDKLRDKLKMQRKPPGFLALVKLLREBEEWEATLGPDRSLEGLVAPRPARI 360
DB 301 ATLPDKLRDKLKMQRKPPGFLALVKLLREBEEWEATLGPDRSLEGLVAPRPARI 360
QY 361 TGVGAVPLPASGNSFDARPSQGVRRRGRGOHRGGVARAGSGRKRKHTFCYSCGD 420
DB 361 TGVGAVPLPASGNSFDARPSQGVRRRGRGOHRGGVARAGSGRKRKHTFCYSCGD 420
```

QY 421 GHIRVQCINPSNLLLVKQKQAAVESGNGNNAWDKSHPKSKAK 463
 DB 421 GHIRVQCINPSNLLLVKQKQAAVESGNGNNAWDKSHPKSKAK 463

RESULT 3

ABB05727
 ID ABB05727 standard; protein; 455 AA.

XX ABB05727;
 XX 30-APR-2002 (first entry)
 XX Human signal transduction protein clone tes3_5k22.
 DE Human; foetal brain; foetal kidney; melanoma; testis; amygdala;
 XX gene therapy.
 XX Homo sapiens.

XX WO2001198454-A2.
 XX 27-DEC-2001.
 XX 25-APR-2001; 2001WO-IB002050.
 XX 25-APR-2000; 2000US-0199380P.

XX (GSHU-) GERMAN HUMAN GENOME PROJECT.
 XX Wiemann S;
 XX WPI; 2002-055860/07.
 XX N-PSDB; ABA93764.
 XX Human cDNA sequences and clones derived from human fetal brain, fetal
 kidney, melanoma, testis and amygdala cDNA libraries, useful in genetic
 screening and therapy.

XX Claim 1; Page 382; 61pp; English.
 XX The present invention describes assemblages and computer readable media
 comprising novel human cDNA sequences and clones derived from human
 foetal brain, foetal kidney, melanoma, testis and amygdala cDNA
 libraries. ABA93702 to ABA93766 represent human cDNA sequences from the
 present invention which encode the proteins given in AB805662 to
 AB805729. The human cDNA sequences and clones can be used in gene
 therapy. The clones may be used in a variety of applications, for example
 they may be used in profiling assays, for providing large arrays of human
 genetic material for implementing large-scale screening strategies and
 for treating diseases via gene therapy procedures

XX Sequence 455 AA;
 Query Match 94.2%; Score 2283.5; DB 5; Length 455;
 Best Local Similarity 98.2%; Pred. No. 3.8e-217;
 Matches 439; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

QY 1 MPTLLQDWCRCGEHLNTRRCMLILGIPEDCGEDEFEEETLQACRHLGRYVIGRMFRREE 60
 DB 1 MPTLLQDWCRCGEHLNTRRCMLILGIPEDCGEDEFEEETLQACRHLGRYVIGRMFRREE 60
 QY 61 NQAAILLELAQIDYALLPRIPKGGPWVIVKPRNSDGEFLNRLNRFLEERTVSDM 120
 DB 61 NQAAILLELAQIDYALLPRIPKGGPWVIVKPRNSDGEFLNRLNRFLEERTVSDM 120
 QY 121 NRVLGSDTNCAPRTVISPEFTWAQTLGRAVQPLLEQMLYRELRFVSGNTISIPGALAF 180
 DB 121 NRVLGSDTNCAPRTVISPEFTWAQTLGRAVQPLLEQMLYRELRFVSGNTISIPGALAF 180
 QY 181 DAWLEHTTEMLOMQVPEGEKRRRIMECLRGFPALQVVSGLFASNASITVECLAAQQVF 240
 DB 181 DAWLEHTTEMLOMQVPEGEKRRRIMECLRGFPALQVVSGLFASNASITVECLAAQQVF 240

QY 241 GPVESHKIAQVKLCKAYQAGEKVSFVLRLEPLLQRAVENNVSRNNVQTRLKRVLSG 300
 DB 241 GPVESHKIAQVKLCKAYQAGEKVSFVLRLEPLLQRAVENNVSRNNVQTRLKRVLSG 300
 QY 301 ATLPDKLRDLKLMKQRRKPPGFALVKLDRREEWEATLGPDRSLRGLEVA PRPPARI 360
 DB 301 ATLPDKLRDLKLMKQRRKPPGFALVKLDRREEWEATLGPDRSLRGLEVA PRPPARI 360
 QY 361 TGVGAVPLPASGNSFDARPSQGYRRRRGRGQHRGGVARAGSRGSRKRKRTHTFCYSCGD 420
 DB 361 TGVGAVPLPASGNSFDARPSQGYRRRRGRGQHRGGVARAGSRGSRKRKRTHTFCYSCGD 420
 QY 421 GHIRVQCINPSNLLLVKQKQAAVESG 447
 DB 421 GHIRVQCINPSNLLLVKQKQAAVESG 446

RESULT 4

AAO16179
 ID AAO16179 standard; protein; 452 AA.

XX AAO16179;
 XX 28-MAR-2003 (first entry)
 XX Human protein #5.
 DE Human; vaccine; adult whole brain; foetal whole brain; tonsil;
 XX adult hippocampus; disease-associated SNP analysis; knockout mouse;
 XX disease model mouse; cancer; neurological disorder.

XX Homo sapiens.
 XX WO200299103-A1.
 XX 12-DEC-2002.

XX 27-MAY-2002; 2002WO-JP005134.
 XX 04-JUN-2001; 2001JP-00168370.
 XX 16-AUG-2001; 2001JP-00246915.

XX (KAZU-) KAZUSA DNA RES INST FOUND.
 XX (PROT-) PROTEIN EXPRESS CO LTD.

XX Ohara O, Nagase T, Nakajima D;
 XX WPI; 2003-140622/13.
 XX N-PSDB; AAL51207.

XX DNA preferentially expressed in human adult and fetal brain tissue useful
 for diagnosis, treatment and analysis of cancer and mental disorders.

XX Claim 1; Page 56-60; 73pp; Japanese.

XX The invention comprises the amino acid and coding sequences of seven
 human proteins that are preferentially expressed in adult whole brain,
 CC foetal whole brain, tonsil and adult hippocampus tissue. The DNA
 CC sequences are useful for the analysis of disease-associated single
 CC nucleotide polymorphisms and the production of knockout and human disease
 CC model mice. The DNA and protein sequences of the invention are useful for
 CC the prevention (vaccine) and treatment of cancer and neurological
 CC disorders. The present amino acid sequence represents a human protein of
 CC the invention

XX Sequence 452 AA;

Query Match 36.5%; Score 883.5; DB 6; Length 452;
 Best Local Similarity 42.6%; Pred. No. 2.8e-78;
 Matches 201; Conservative 75; Mismatches 155; Indels 41; Gaps 10;

QY 1 MPTLLQDWCRCGEHLNTRRCMLILGIPEDCGEDEFEEETLQACRHLGRYVIGRMFRREE 60

[illegible]

Db	121	FRALQGEVSPATVPCISPELLAHLLQAMAHAPQPLL-PMRYKRLRVFSGSAVPAPBEE	179
Qy	179	AFDAWLEHTTEMLQMWQVPEGEKRRRLMECLRGPALQVVSGLSRASNASITVEECIALAQ	238
Db	180	SFEWLEQATEIYKWPVTAEEKRWLAESLRGPAIDLMHVQADNPISIVEECLEAFKQ	239
Qy	239	VFGVESHKIAQVKLCKAYQAGEKVSFVLRLPELLQRAVENNVSRNNVQNTLKRVL	298
Db	240	VFGLESRRQTQVRYLTKYQEEGEKVSAYVLRLTLLRRRAVEKRAIPRIADOVLEQVM	299
Qy	299	SGATLPDKLRDKLKLMKQRKPPGCFIALVKLLRREEEWEATLGDRESLE	348
Db	300	AGATLNMICWLRRLKDCGPPSPFLMKVIRSEEEAGF--ENEGIE	347

patients. The Ma2 gene and its transcription and/or translation products may have a causative role in the regional selective neuronal degeneration typically observed in AD, or may confer a neuroprotective function to the remaining nerve cells. Methods are claimed for diagnosing or prognosticating a neurodegenerative disease, for monitoring the progression of a neurodegenerative disease, and for evaluating treatment of a neurodegenerative disease, especially AD, in a subject by determining the level and/or activity of a transcription or translation product of an Ma onconeural antigen gene, especially Ma2. Also claimed are: a method for treating or preventing AD and related neurodegenerative disorders using the Ma2 gene or its transcription or translation product; a method of screening for modulating agents of neurodegenerative diseases; and a recombinant non-human animal comprising an Ma2 gene sequence, which is useful for screening, testing and validating candidate diagnostic and therapeutic agents.

RESULT 6	
AA874701	
ID	AA874701 standard; protein; 353 AA.
XX	
XX	AA874701;
XX	
XX	12-JUN-2001 (first entry)
XX	
XX	Human membrane associated protein MEMAP-7.
XX	
XX	Human; membrane associated protein; MEMAP; diagnosis; cytostatic;
KW	antiinflammatory; anticonvulsant; immunosuppressive; anti diarrheic;
KW	arteriosclerotic; gene therapy; cell proliferative disorder;
KW	autoimmune disorder; inflammatory disorder; neurological disorder;
KW	gastrointestinal disorder; cancer; inflammation; atherosclerosis;
KW	epilepsy; diarrhoea.
XX	
XX	Homo sapiens.
OS	
XX	
XX	WO200112662-A2.
PN	
XX	
XX	22-FEB-2001.
PD	
XX	
XX	14-AUG-2000; 2000WO-US022315.
PF	
XX	
XX	17-AUG-1999; 99US-0149641P.
PR	

```
PR 09-NOV-1999; 99US-0164203P.
XX (INCY-) INCYTE GENOMICS INC.
XX Lal P, Yue H, Tang YT, Bandman O, Burford N, Azimzai Y;
PI Baughn MR, Lu DAM, Patterson C;
XX WPI; 2001-168860/17.
XX N-PSDB; AAF81747.
XX Isolated polypeptide with a human membrane associated protein sequence is
XX useful for the diagnosis, prevention and treatment of cell proliferative,
XX autoimmune/inflammatory, neurological and gastrointestinal disorders.
XX Claim 1; Page 119-120; 173pp; English.
XX AAF81741 to AAF81777 encode the human membrane associated proteins
XX (MEMAP) given in AAB74695 to AAB74731. MEMAPs have cytostatic,
XX antiinflammatory, anticonvulsant, immunosuppressive, antidiarrheic and
XX antiarteriosclerotic activities, which can be used in gene therapy.
XX MEMAPs and agonist of MEMAPs can be used to treat a disease or condition
XX associated with decreased expression of functional MEMAP and antagonists
XX of MEMAP are used to treat a disease or condition associated with
XX overexpression of functional MEMAP. These disorders include cell
XX proliferative, autoimmune/inflammatory, neurological and gastrointestinal
XX disorders. The MEMAP polynucleotides and proteins are also used for the
XX diagnosis of these disorders. Specific examples of these disorders
XX include cancer, inflammation, atherosclerosis, epilepsy and diarrhoea.
XX MEMAP proteins can be used to screen for compounds which specifically
XX bind MEMAP including antibodies, oligonucleotides, proteins and small
XX molecules. MEMAP polynucleotides can be used to prepare transgenic
XX animals which can be studied to provide information concerning human
XX disease. Anti-MEMAP antibodies are useful in immunoassays for the
XX detection of MEMAP protein and can be used as antagonists to treat or
XX prevent a disorder associated with MEMAP. Polynucleotides encoding MEMAP
XX can be delivered to target cells with genetic abnormalities with respect
XX to the expression of MEMAP to treat or prevent a disorder associated with
XX MEMAP
XX
XX Sequence 353 AA;
Query Match 33.8%; Score 818; DB 4; Length 353;
Best Local Similarity 43.6%; Pred. No. 6e-72;
Matches 172; Conservative 52; Mismatches 113; Indels 10; Gaps 4;
Qy 1 MPFLQLQDWCGRGHLNTRRCMLILGIPEDCCGEFEETLQACRHLGRYRIGRMFRREE 60
Db 1 MAMTLLEDGCGMDVNSQALLVWGIPVNCDEAEIETLQAAQPVS-YRMLGRMFWRREE 59
Qy 61 NQAIIILEAQDIDYALLPREIPQKGPWEVIVKPRNSDGFENLRNLFLEERRTVSDM 120
Db 60 NAKAALLELTGAVDYAAIPREMPGKGGMVKVLPKPTSDAEFLRLHLFLAREGWTQDV 119
Qy 121 NRVLGSDTNCSPRVTTISPEWTTAQTLLGAQVQLLEOMLYRELRFVSGNTISIPGALAF 180
Db 120 ARVLGFQNPPTTPGEPMAELNY--ILDNVIQFLVESIWYKRLTLFSGRDIPOGGETF 177
Qy 181 DAWLEHTTEMLQWQVEGEXRRRLMECLRGFALQVWSGLRASNASTIVVEECLAALQOVF 240
Db 178 DPWLEHTNEVLVEEQVSDVEKRRRLMESLRGPAADVIRILKSNPNPAITTAECLEKALEQVF 237
Qy 241 GPVESHKIAQVKLCKAQAEAGEKVSFVLRLEPLLQRAVENNVSRNNVQTRKRVLSG 300
Db 238 GSVESSRDAQIKFLNTYQNPGEKLSAYVIRLEPLLQKVEKGKADKDNVNOARLEQVIAG 297
Qy 301 ATLDPKLRDLKLMKQRKRP-PGFLALVKLLR-----EESEWEATL 340
Db 298 ANHSGAIRRQLWLTGAGSGPAPNLFQLLVQVIRREAEKEEBEAEATL 344
RESULT 7
AAE01340
ID AAE01340 standard; protein; 351 AA.
```

```
XX AAE01340;
XX 17-JUL-2001 (first entry)
XX Human gene 22 encoded secreted protein fragment, SEQ ID NO:205.
XX Human; secreted protein; proliferative disorder; cancer; tumour;
XX foetal abnormality; developmental abnormality; haematopoietic disorder;
XX immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
XX inflammation; allergy; neurological disorder; Alzheimer's disease;
XX Parkinson's disease; cognitive disorder; schizophrenia; asthma;
XX skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
XX cardiovascular disorder; angiogenic disorder; kidney disorder;
XX gastrointestinal disorder; pregnancy-related disorder;
XX endocrine disorder; infection; wound healing; vulnerability; cell culture;
XX chemotaxis; food additive; gene therapy; binding partner identification.
XX Homo sapiens.
XX WO200134769-A2.
XX 17-MAY-2001.
XX 01-NOV-2000; 2000WO-US030040.
XX 05-NOV-1999; 99US-0163580P.
XX 30-JUN-2000; 2000US-0215130P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Ruben SM, Komatsoulis GA, Wei P, Baker KP, Fiscella M;
XX WPI; 2001-308781/32.
XX New isolated nucleic acid molecule encoding a human secreted protein is
XX used in preventing, treating or ameliorating a medical condition.
XX Disclosure; Page 46; 519pp; English.
XX AAD05121-AA050203 represent cDNAs corresponding to 24 human secreted
XX protein genes, and AAE01232-AAE01311 represent the proteins they encode.
XX AAE01312-AAE01340 represent human secreted protein variants or fragments.
XX The secreted proteins and their genes are useful for preventing, treating
XX or ameliorating medical conditions, e.g., by protein or gene therapy.
XX Pathological conditions can be diagnosed by determining the amount of the
XX new protein in a sample or by determining the presence of mutations in
XX the new genes. Specific uses are described for each of the 24 genes,
XX based on the tissues in which they are most highly expressed, and include
XX developing products for the diagnosis or treatment of proliferative
XX disorders, cancer, tumours, foetal and developmental abnormalities,
XX haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
XX diseases (e.g., rheumatoid arthritis), inflammation, allergies,
XX neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
XX cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
XX psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
XX angiogenic disorders, kidney disorders, gastrointestinal disorders,
XX pregnancy-related disorders, endocrine disorders, and infections. The
XX proteins can also be used to aid wound healing and epithelial cell
XX proliferation, to prevent skin aging due to sunburn, to maintain organs
XX before transplantation, for supporting cell culture of primary tissues,
XX to regenerate tissues, to identify their cognate ligands or binding
XX partners, and in chemotaxis, and can be used as a food additive or
XX preservative to modify storage properties. Antibodies specific for a
XX protein of the invention can be used in alleviating symptoms associated
XX with the disorders mentioned above, and in diagnostic immunoassays e.g.,
XX radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The
XX present sequence represents a human secreted protein fragment referred to
XX in the disclosure of the invention
XX Sequence 351 AA;
Query Match 31.7%; Score 768.5; DB 4; Length 351;
```

Best Local Similarity 46.6%; Pred. No. 4.8e-67;
Matches 165; Conservative 60; Mismatches 118; Indels 11; Gaps 6;

QY 1 MPTLLQDRCRGEHLNTRRCMLILGIPEDCGDEFEETLQACRHLGRVIRGMPFRREE 60
Db 1 MTLRLLEDRCRGMNPRKALLIAGISQSCSVAEIEALQAGLAPLGEYLLGRMFRDE 60

QY 61 NQAAILLELAQDIDYALLPREIPKGGPWEVIVKPRNSDGEFLNRLNRFLEERRTVSDM 120
Db 61 NRKVALVGLTAETSHALVPEIKPGKGIWRVIFKPPDPNTFLSRNLEFLAGEGTMVGL 120

QY 121 NRVLGSDTNCSPRVITISPEFWT--WAQTLGAAVQPLLEQMLYRELRFVSGNTISIPGAL 178
Db 121 SRALGHENGLDPEQGMIPEMWAPMLAQAAL-EALQPALQCLKYKLVFSGRESPEEGE 179

QY 179 AFDWLEHTTEMLQWQVPEGEKRRLEMECLRGALQVVGSLRASNASITVECLAAQQ 238
Db 180 EFRGMFHTTQMIKAWQVPEVEKRRLLSRLGPDALDVIRVLKINNPLITVDECLQALBE 239

QY 239 VFGPVESHKIAQVLCCKAYQAGEKVSFVLRLPLELQRAVENNVVSRNVTQLRKYVL 298
Db 240 VFGVTDNPRELQVKYLTYYQKDEKLSAYVLRLEPLLQKLVQGAIERDAVNAQRLDQVI 299

QY 299 SGATLPDKLRDKLKMQRKPPGFALVKKLRE---EEWEATLGPDRSLEG 349
Db 300 AGA-VHKTIRRELNL-PEDGPAPGFLQLLVLIKDYEAEEEEALL---QAILEG 348

RESULT 8
AAB74695
ID AAB74695 standard; protein; 351 AA.
AC AAB74695;
XX
XX
DT 12-JUN-2001 (first entry)
DE Human membrane associated protein MEMAP-1.
XX
KW Human; membrane associated protein; MEMAP; diagnosis; cytostatic;
KW antiinflammatory; anticonvulsant; immunosuppressive; antidiarrheic;
KW antiarteriosclerotic; gene therapy; cell proliferative disorder;
KW autoimmune disorder; inflammatory disorder; neurological disorder;
KW gastrointestinal disorder; cancer; inflammation; atherosclerosis;
KW epilepsy; diarrhoea.
XX
OS Homo sapiens.
XX
XX
PN WO200112662-A2.
XX
PD 22-FEB-2001.
XX
PF 14-AUG-2000; 2000WO-US022315.
XX
PR 17-AUG-1999; 99US-0149641P.
PR 09-NOV-1999; 99US-0164203P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
XX Lal P, Yue H, Tang YT, Bandman O, Burford N, Azimzai Y;
PI Baughn MR, Lu DAM, Patterson C;
XX
XX WPI; 2001-168860/17.
DR N-PSDB; AAF81741.
XX
XX Isolated polypeptide with a human membrane associated protein sequence is
PT useful for the diagnosis, prevention and treatment of cell proliferative,
PT autoimmune/inflammatory, neurological and gastrointestinal disorders.
XX
PS Claim 1; Page 114-115; 173pp; English.
XX
XX AAF81741 to AAF81777 encode the human membrane associated proteins
CC (MEMAP) given in AAB74695 to AAB74731. MEMAPs have cytostatic,
CC antiinflammatory, anticonvulsant, immunosuppressive, antidiarrheic and

CC antiarteriosclerotic activities, which can be used in gene therapy,
CC MEMAPs and agonist of MEMAPs can be used to treat a disease or condition
CC associated with decreased expression of functional MEMAP and antagonists
CC of MEMAP are used to treat a disease or condition associated with
CC overexpression of functional MEMAP. These disorders include cell
CC proliferative, autoimmune/inflammatory, neurological and gastrointestinal
CC disorders. The MEMAP polynucleotides and proteins are also used for the
CC diagnosis of these disorders. Specific examples of these disorders
CC include cancer, inflammation, atherosclerosis, epilepsy and diarrhoea.
CC MEMAP proteins can be used to screen for compounds which specifically
CC bind MEMAP including antibodies, oligonucleotides, proteins and small
CC molecules. MEMAP polynucleotides can be used to prepare transgenic
CC animals which can be studied to provide information concerning human
CC disease. Anti-MEMAP antibodies are useful in immunoassays for the
CC detection of MEMAP protein and can be used as antagonists to treat or
CC prevent a disorder associated with MEMAP. Polynucleotides encoding MEMAP
CC can be delivered to target cells with genetic abnormalities with respect
CC to the expression of MEMAP to treat or prevent a disorder associated with
CC MEMAP
XX
SQ Sequence 351 AA;
Query Match 31.7%; Score 768.5; DB 4; Length 351;
Best Local Similarity 46.6%; Pred. No. 4.8e-67;
Matches 165; Conservative 60; Mismatches 118; Indels 11; Gaps 6;

QY 1 MPTLLQDRCRGEHLNTRRCMLILGIPEDCGDEFEETLQACRHLGRVIRGMPFRREE 60
Db 1 MTLRLLEDRCRGMNPRKALLIAGISQSCSVAEIEALQAGLAPLGEYLLGRMFRDE 60

QY 61 NQAAILLELAQDIDYALLPREIPKGGPWEVIVKPRNSDGEFLNRLNRFLEERRTVSDM 120
Db 61 NRKVALVGLTAETSHALVPEIKPGKGIWRVIFKPPDPNTFLSRNLEFLAGEGTMVGL 120

QY 121 NRVLGSDTNCSPRVITISPEFWT--WAQTLGAAVQPLLEQMLYRELRFVSGNTISIPGAL 178
Db 121 SRALGHENGLDPEQGMIPEMWAPMLAQAAL-EALQPALQCLKYKLVFSGRESPEEGE 179

QY 179 AFDWLEHTTEMLQWQVPEGEKRRLEMECLRGALQVVGSLRASNASITVECLAAQQ 238
Db 180 EFRGMFHTTQMIKAWQVPEVEKRRLLSRLGPDALDVIRVLKINNPLITVDECLQALBE 239

QY 239 VFGPVESHKIAQVLCCKAYQAGEKVSFVLRLPLELQRAVENNVVSRNVTQLRKYVL 298
Db 240 VFGVTDNPRELQVKYLTYYQKDEKLSAYVLRLEPLLQKLVQGAIERDAVNAQRLDQVI 299

QY 299 SGATLPDKLRDKLKMQRKPPGFALVKKLRE---EEWEATLGPDRSLEG 349
Db 300 AGA-VHKTIRRELNL-PEDGPAPGFLQLLVLIKDYEAEEEEALL---QAILEG 348

RESULT 9
AAU08664
ID AAU08664 standard; protein; 351 AA.
XX
AC AAU08664;
XX
XX 18-DEC-2001 (first entry)
XX
XX Human NOV8 protein.
DE
XX
KW Human; NOV8; cytostatic; neuroprotective; vulnerary;
KW cerebroprotective; antiparkinsonian; hypotensive; antiasthmatic;
KW antidiabetic; antipsoriatic; antiinflammatory; immunosuppressive;
KW antiatherosclerotic; dermatological; cancer; neurological disorder;
KW Parkinson's disease; diabetes mellitus; asthma; enamel defect;
KW immune disorder; autoimmune disease; respiratory disorder; bone disorder;
KW musculoskeletal disorders; leukaemia; lymphoma;
KW cell growth regulation disorder; lesional psoriatic skin;
KW atherosclerosis; abdominal aortic aneurysm.
XX
OS Homo sapiens.
XX

PN WO200169851-A2.
 XX 20-SEP-2001.
 XX 12-MAR-2001; 2001WO-US007735.
 XX 10-MAR-2000; 2000US-0188277P.
 XX 10-MAR-2000; 2000US-0188316P.
 XX 14-MAR-2000; 2000US-0189139P.
 XX 14-MAR-2000; 2000US-0189140P.
 XX 17-MAR-2000; 2000US-0190231P.
 XX 17-MAR-2000; 2000US-0190401P.
 XX (CURA-) CURAGEN CORP.
 XX Padigar M, Vernet CAM, Fernandes E, Shimkets RA, Spaderna SK;
 PI Majumder K, Li L;
 XX DR N-PSDB; AAS13342.
 XX WPI; 2001-570869/64.
 XX N-PSDB; AAS13342.
 XX Novel polypeptides and nucleic acids homologous to members of collagen,
 PT potassium channel, tufelin family of proteins for diagnosing, treating
 PT cancer, atherosclerosis, neurological, skin and enamel defect disorders.
 XX
 XX Claim 1; Page 29; 128pp; English.
 XX The invention relates to isolated NOVX (NOVX1-11) polypeptides and the
 CC polynucleotides that encode them. NOVX polypeptides, polynucleotides and
 CC anti-NOVX antibodies are useful for treating or preventing a pathology
 CC associated with NOVX polypeptide in humans and for treating a syndrome
 CC associated with human disease e.g. disorders characterised by altered
 CC cell motility, proliferation and migration e.g. cancer, angiogenesis and
 CC wound healing (NOV1-3), neurological disorders, e.g. episodic ataxia,
 CC autosomal dominant myokymia, stroke, Parkinson's disease, Alzheimer's
 CC disease, non-insulin dependent diabetes mellitus, asthma, hypertension
 CC and seizure (NOV4), enamel defects, such as amelogenesis imperfecta and
 CC disorders involving enamel defects, including hypoplasia and
 CC hypomineralisation (NOV5-7), paraneoplastic neurological disorders, e.g.
 CC paraneoplastic limbic of brain-stem encephalitis occurring during
 CC testicular cancer, diabetes, reproductive health, metabolic and endocrine
 CC disorders, gastrointestinal disorders, immune disorders and autoimmune
 CC diseases, respiratory disorders, bone disorders, musculoskeletal
 CC disorders, leukaemia/lymphoma and tissue/cell growth regulation disorders
 CC (NOV8), lesional psoriatic skin (NOV9-10) and atherosclerosis, abdominal
 CC aortic aneurysm and neurological disorders (NOV11). NOVX polypeptide is
 CC also useful for identifying an agent that binds to it and a cell
 CC expressing NOVX polypeptide is useful for identifying a therapeutic agent
 CC for use in treatment of a NOVX related pathology. The antibodies and a
 CC polypeptide having 95% sequence identity to NOVX polypeptide are useful
 CC for treating a pathological state in a mammal. The present sequence
 CC represents NOV8, a possible neuronal antigen-like protein
 XX
 XX Sequence 351 AA;
 Query Match 31.7%; Score 768.5; DB 4; Length 351;
 Best Local Similarity 46.8%; Pred. No. 4.8e-67;
 Matches 165; Conservative 60; Mismatches 118; Indels 11; Gaps 6;
 QY 1 MPTLLQDWCRCGHEHNLTRCMLILGIPEDCGDEFEETLQEAACRHILGRYRVIGRMFRREE 60
 DB 1 MTLRLLEDWCRCGMDMNPRAKLLIAGISQSCVAIEEALQAGLAPLGEVLLGRMFRDE 60
 QY 61 NQAAILLELAQDIDYALLPREIPGKGGPWEIVKPRNSDGEFEETLQEAACRHILGRYRVIGRMFRREE 120
 DB 61 NRKVALVGLTAETSHALVPKEIPKGGIWRVIFKPPDPDNTFLSLRNEFLAGEGMTGEL 120
 QY 121 NRVLGSDTNCSPARVTISPEFWT--NAQTILGAQVPLLEQMLYRELVRVPSGNTISIPGAL 178
 DB 121 SRAIGHENGSLDPEQGMIPENWAPMLAQAAL-EALQPALQCLKYKILRVFSGRESPEEGEE 179
 QY 179 AFDWLHHTTTEMLQMWQVPEGEKRRRLMEICRGALPQVYSGLRASNASTIVBECLAAALQ 238

Db 121 SVAGALG-----VGLRRV-----CWLRSIGAVQVPWVEAVRCQSLGVTSGRDQAPGEE 169
QY 179 AFDWLHTTEMLQMWQ-VPEGEKRRRLMECLGPGALQVVGSLGRASNASTIVVECLAAIQ 237
Db 170 SPEVLDHTTEMLHVWQGVSERRRRRLLEGGRTALQVLVHALLAENPARTAQDCUAAIA 229
QY 238 QVFGPVESHKIAQVKLCCKAYQAGEKVVSVFLRLEPLLQRAVENNVVSRNNVQTRKXV 297
Db 230 QVFGDNESQATIRVKCLTAQQQSGERLSAFVLEVLVLLQKAMEKEALASADRVLQRM 289
QY 298 LSGATLPDKLRDKLKMQRKPPGFLALVKLLRBESEWEATLG----PDRESLEGLEVA 353
Db 290 LTRAHLTEPLDALRKLRMAGRSFSLMLGLVRESEAWESLARSVRAQTQEGAGARAG 349
QY 354 PRPPARI-TGVGAVP-----LPASGNSFDARPSQGYR 384
Db 350 AQAVARASTKVEAVPGGPREPEGLLQAGQAEELLQEGLK 391

RESULT 11

ABG99947
ID ABG99947 standard; protein; 399 AA.
XX
AC ABG99947;
XX
DT 17-JAN-2003 (first entry)
XX
DE Human novel polypeptide #60.
XX
KW Human; genetic disorder; gene mapping; medical imaging; cancer;
KW neurodegenerative disorder; lymphoid cell disorder; osteoporosis;
KW Parkinson's disease; Alzheimer's disease; bone degenerative disorder;
KW osteoarthritis; periodontal disease; liver fibrosis; viral infection;
KW fungal infection; bacterial infection; autoimmune disease; diabetes;
KW atopic dermatitis.
XX
OS Homo sapiens.
XX
FN WO200274961-A1.
XX
PD 26-SEP-2002.
XX
PF 14-MAR-2002; 2002WO-US005109.
XX
PR 15-MAR-2001; 2001US-00810173.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Zhou P, Goodrich R, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
PI Wehrman T, Wang J, Wang D, Drmanac RT;
XX
DR WPI; 2003-040556/03.
DR N-PSDB; ABX05045.
XX
PT New isolated polypeptides and polynucleotides, useful for preventing,
PT treating or ameliorating medical conditions, such as cancer,
PT neurodegenerative disorders, lymphoid cell disorders, bone degenerative
PT disorders, and infections.
XX
PS Claim 9; SEQ ID NO 586; 235pp; English.

XX
CC The invention relates to human polynucleotides and the polypeptides they
CC encode. The polynucleotides and polypeptides are useful in diagnostics,
CC forensics, gene mapping, medical imaging, identification of mutations,
CC responsible for genetic disorders or other traits, assessing biodiversity
CC and producing many other types of data and products dependent on DNA and
CC amino acid sequences. They are also useful for preventing, treating or
CC ameliorating medical conditions, such as cancer, neurodegenerative
CC disorders (e.g. Parkinson's disease, Alzheimer's disease), lymphoid cell
CC disorders, osteoporosis, osteoarthritis, bone degenerative disorders,
CC periodontal disease, liver fibrosis, infections (e.g. viral, fungal or

CC bacterial) or autoimmune diseases (e.g. diabetes, atopic dermatitis).
CC Sequences ABG9988-ABG9989 and ABU0010-ABU00433 represent human
CC polypeptides of the invention. Note: the sequence data for this patent is
CC not represented in the printed specification but is based on sequence
CC information supplied by the European Patent Office
XX
SQ Sequence 399 AA;

Query Match 31.7%; Score 768.5; DB 6; Length 399;
Best Local Similarity 42.5%; Pred. No. 5.9e-67;
Matches 171; Conservative 72; Mismatches 130; Indels 29; Gaps 7;

QY 1 MPTLLQDWCRGHLNTRRCMLILGIPEDCGDEPETLQEAQRHLYRVRVIGRMFRREE 60
Db 1 MAVTWLQDWCRWGVNARRGLLILGIPEDCDAAEQESLEAALRPMGHFTVLGKAREED 60
QY 61 NQAQILLELAQDIDYALLPREIPGKGPWEVIVKPRNSDGEF--LNRLNRFLEEEERTVS 118
Db 61 NATAALVELDREVNVALVPREIPGTGGPWNVVFPRCSGEEFLGLGRVFFHPEQGMVBE 120
QY 119 DMNRVLGSDTNCASAPRVITISPEFTWQAOTLGAQVQPLEQMLYRELRFVSGNTISIPGAL 178
Db 121 SVAGALG-----VGLRRV-----CWLRSIGAVQVPWVEAVRCQSLGVTSGRDQAPGEE 169
QY 179 AFDWLHTTEMLQMWQ-VPEGEKRRRLMECLGPGALQVVGSLGRASNASTIVVECLAAIQ 237
Db 170 SPEVLDHTTEMLHVWQGVSERRRRRLLEGGRTALQVLVHALLAENPARTAQDCUAAIA 229
QY 238 QVFGPVESHKIAQVKLCCKAYQAGEKVVSVFLRLEPLLQRAVENNVVSRNNVQTRKXV 297
Db 230 QVFGDNESQATIRVKCLTAQQQSGERLSAFVLEVLVLLQKAMEKEALASADRVLQRM 289
QY 298 LSGATLPDKLRDKLKMQRKPPGFLALVKLLRBESEWEATLG----PDRESLEGLEVA 353
Db 290 LTRAHLTEPLDALRKLRMAGRSFSLMLGLVRESEAWESLARSVRAQTQEGAGARAG 349
QY 354 PRPPARI-TGVGAVP-----LPASGNSFDARPSQGYR 384
Db 350 AQAVARASTKVEAVPGGPREPEGLLQAGQAEELLQEGLK 391

RESULT 12

AAAB12525
ID AAAB12525 standard; protein; 329 AA.
XX
AC AAAB12525;
XX
DT 02-NOV-2000 (first entry)
XX
DE Human Mal protein SEQ ID NO:4.
XX
KW Mal1, Mal2; Mal3; Mal4; Mal5; paraneoplastic limbic encephalitis; diagnosis;
KW paraneoplastic syndrome; neoplasm; colon cancer; brainstem encephalitis;
KW breast cancer; parotid gland cancer; lung cancer; testicular cancer;
KW germ-cell tumour.
XX
OS Homo sapiens.
XX
FN JP2000146982-A.
XX
PD 26-MAY-2000.
XX
PF 10-NOV-1999; 99JP-00320171.
XX
PR 10-NOV-1998; 98US-00189527.
XX
PA (SLOK) SLOAN KETTERING INST CANCER RES.
XX
DR WPI; 2000-468119/41.
XX N-PSDB; AAA60833.
XX
PT Diagnosing paraneoplastic syndrome e.g. paraneoplastic limbic
PT encephalitis or neoplasm e.g. colon cancer comprising assessing a test

PT sample for the presence or absence of antibodies to a Ma family

PT polypeptide.

XX Claim 48; Fig 1; 27pp; Japanese.

XX

XX The present invention describes a method for diagnosing a paraneoplastic
CC syndrome or neoplasm. The method comprises assessing a test sample for
CC the presence or absence of antibodies to a Ma family polypeptide (I). The
CC method is used to diagnose a paraneoplastic syndrome especially
CC paraneoplastic limbic encephalitis and/or brainstem encephalitis or
CC neoplasm in an individual. The method diagnoses the neoplasm by assessing
CC antibodies to (I) preferably Ma1, which is indicative presence of breast
CC cancer, colon cancer, parotid gland cancer, lung cancer, testicular
CC cancer, and germ-cell tumours or Ma2, which is indicative of testicular
CC cancer, germ-cell tumour, and lung cancer. The present sequence is the
CC Ma1 protein as given in the present invention

XX Sequence 329 AA;

Query Match 31.2%; Score 755.5; DB 3; Length 329;

Best Local Similarity 48.8%; Pred. No. 8.5e-66;

Matches 156; Conservative 50; Mismatches 111; Indels 3; Gaps 2;

QY 1 MPTLLQDWCGRGHLNTRRCMLILGIPDCGDEFEETLOEACRHLGRYVIGRMFRRE 60

DB 1 MAMTLLDWCGRGMDVNSQRILLVWGI PVNCDAAEIEETLQAAMPQVS-YRMLGRMFRRE 59

QY 61 NQAAILLELAQIDYALLPREIPKGGPWEVIVKPRNSDGEFLNRLNRFLEERTVSDM 120

DB 60 NAKAAMELGTAGVDYAAIPREMPKGGVWKVLPKPTSDAEFLERLHLLAREGWTQDV 119

QY 121 NRVLGSDTNCAPRVITISPEFWTQAOTLGAQVPLLEQMLYRELRFVSGNTISIPGALAF 180

DB 120 ARVLGFONFTPTPGPMPAEMLYN-ILDNVIOPLVESIWKYRLTFSFGHPRAWGNF 177

QY 181 DAMLEHTTEMLQMWQVPEGEKRRRLMECLRGPAQVVSGLRASNASITVECLAAQOVF 240

DB 178 DPWLEHNEVLEWQVSDVEKTRLMESLRGPAADVIRLKSNPAITAECLALEHVP 237

QY 241 GPVESHKIAQVKLCKAYOAGEKVSFVLRLEPLQLQRAVENNVSRNNVQTRKRVLSG 300

DB 238 GSVESSRDAQIKFLNTYQNPGEKLSAVVIRLEPLQKVKGEAIDKDNVQARLEQVIAG 297

QY 301 ATLPDLKRDKLKMKQRKP 320

DB 298 ANHSGAIRQLWLGTAREGP 317

RESULT 13

AAB43023

ID AAB43023 standard; protein; 312 AA.

XX AAB43023;

AC

XX 08-FEB-2001 (first entry)

DT

XX

DE Human ORFX ORF2787 polypeptide sequence SEQ ID NO:5574.

XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnary; antipsoriatic; antiparkinsonian; nontropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disease; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive.

OS Homo sapiens.

XX WO200058473-A2.

XX 05-OCT-2000.

XX 31-MAR-2000; 2000WO-US008621.

XX 31-MAR-1999; 99US-0127607P.

XX 02-APR-1999; 99US-0127636P.

XX 05-APR-1999; 99US-0127728P.

XX 30-MAR-2000; 2000US-00540763.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach M;

XX MPI; 2000-602362/57.

XX N-PSDB; AAC77232.

XX Novel nucleic acids and peptides derived from open reading frame X,

XX useful for treating e.g. cancers, proliferative disorders,

XX neurodegenerative disorders and cardiovascular disease.

XX Claim 11; Page 4759; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,

XX which represent the human ORFX open reading frames 1 to 3161. The ORFX

XX sequences have activities such as: cytostatic; hepatotropic; vulnary;

XX antipsoriatic; antiparkinsonian; nontropic; neuroprotective; osteopathic;

XX anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;

XX cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;

XX dermatological; immunosuppressive; antiinflammatory; antibacterial;

XX antiviral; antifungal; antirheumatic; antithyroid; and antianemic. The

XX sequences can be used for determining the presence of or predisposition

XX to, or preventing or treating pathological conditions associated with an

XX ORFX-associated disorder. The nucleic acids can be used to express ORFX

XX proteins in gene therapy vectors. The proteins and nucleic acids may be

XX used to treat cancers, proliferative disorders, neurodegenerative

XX disorders, osteoarthritis, graft vs host disease, cardiovascular disease,

XX diabetes mellitus, hypertension, hypothyroidism, cholesterol ester

XX storage, systemic lupus erythematosus, severe combined immunodeficiency

XX (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune

XX disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and

XX cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to

XX enhance coagulation; to inhibit thrombosis; and as a contraceptive

XX Sequence 312 AA;

Query Match 30.6%; Score 740.5; DB 3; Length 312;

Best Local Similarity 48.8%; Pred. No. 2.4e-64;

Matches 148; Conservative 51; Mismatches 101; Indels 3; Gaps 2;

QY 1 MPTLLQDWCGRGHLNTRRCMLILGIPDCGDEFEETLOEACRHLGRYVIGRMFRRE 60

DB 1 MTLRLLEDWCGRGMDMNPCKALLIAGTISQSCSVAIEEALQAGLAPLGEYLLGRMFRDE 60

QY 61 NQAAILLELAQIDYALLPREIPKGGPWEVIVKPRNSDGEFLNRLNRFLEERTVSDM 120

DB 61 NRKVALVGLTAETSHALVPKEIPKGGIWRVIFKPPDPDNTFLSRLNEFLAGEGTMVGL 120

QY 121 NRVLGSDTNCAPRVITISPEFWT--WAOTLGAQVPLLEQMLYRELRFVSGNTISIPGAL 178

DB 121 SRALGHENGSLDPEQGMIPENWAPMLAQAAL-EALQALQCLKYKVKLVFSGRSPPEGEE 179

QY 179 AFDALWHTTEMLQMWQVPEGEKRRRLMECLRGPAQVVSGLRASNASITVECLAAQO 238

DB 180 EFGWPHHTTQMLNAQVDPVVERRRRLLESRLGPDVIRVLKINPLIITVECLQALEE 239

QY 239 VFGVESHKIAQVKLCKAYOAGEKVSFVLRLEPLQLQRAVENNVSRNNVQTRKRVLSG 298

DB 240 VFGVTDNPRELQVKYLLTTYQKDEKLSAYVLRLEPLQLQVQGAERDAVNQARLDQVI 299

CC partners, and in chemotaxis, and can be used as a food additive or
CC preservative to modify storage properties. Antibodies specific for a
CC protein of the invention can be used in alleviating symptoms associated
CC with the disorders mentioned above, and in diagnostic immunoassays e.g.,
CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The
CC present sequence represents a human secreted protein fragment referred to
CC in the disclosure of the invention
XX
SQ Sequence 280 AA;

Query Match 19.7%; Score 476.5; DB 4; Length 280;
Best Local Similarity 41.1%; Pred. No. 3.2e-38;
Matches 117; Conservative 48; Mismatches 97; Indels 23; Gaps 8;
Qy 70 AODIDYALLPREIFKGGPWEVIVKPNNSDGEFLNRLNRFLEEERRRTVSDMNRVLGSDTN 129
Db 11 ASQVQAVLLPQ-----GSRVAVSRARAEFG-----TRFLAGEGMTVGELSRALGHENG 58
Qy 130 CSAPRVTISPEFWT--WAOTLGAAVOPLLEOMLYRELVRVFSGNTISIPGALAFDAWLEHT 187
Db 59 SLDPEQQWPEWAPMLAQAL-EALQPALQCLKYKURVFSGRSPGPEEEFGRWMPHT 117
Qy 188 TEMPLQMVPPEGEKRRIMECLRGPALQVVSGLRASNASITVBECLAAALQOVFGPVESHK 247
Db 118 TQMIKAMQVPDVEKRRLLSIRGPDLDVIRVLKINNPLITVDECLQALEEVFGVTNPR 177
Qy 248 IAQVKLCAYOAGEKVSFVIRLEPLLQRAVENNVSRNNVNOTRLKRVLSGATLPDKL 307
Db 178 ELQVKYLTYYQKDEKLSAYVIRLEPLLQKLVORGAIERDAVNOARLDQVIAGA-VHKTI 236
Qy 308 RDKLKLKMKQREKPPGFLALVKLLRE---EEWEATLGPDRSLEG 349
Db 237 RREUNL-PEDGPAFGFLQLLVILINDYEAEEEEALL---QAILEG 277

Search completed: September 27, 2004, 17:08:24
Job time : 79.9232 secs

Blank

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 27, 2004, 17:05:06 ; Search time 18.7467 Seconds
(without alignments)
2507.755 Million cell updates/sec

Title: US-10-037-860-9
Perfect score: 766
Sequence: 1 DLHVIQADNPISVBECLF.....SIEPERDGYGRWNHGGDD 149

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues 1017041

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	766	100.0	149	4	Q9UL43	Q9UL43 homo sapien
2	758	99.0	364	4	Q94959	Q94959 homo sapien
3	755	98.6	283	4	Q9UL42	Q9UL42 homo sapien
4	741	96.7	364	6	Q9GMU3	Q9GMU3 macaca fasc
5	600.5	78.4	365	11	Q8BHK0	Q8BHK0 mus musculus
6	323.5	42.2	466	11	Q8JZW8	Q8JZW8 mus musculus
7	323	42.2	455	4	Q9H0A4	Q9H0A4 homo sapien
8	323	42.2	463	4	Q9UL41	Q9UL41 homo sapien
9	295.5	38.6	192	4	Q8ND90	Q8ND90 homo sapien
10	295.5	38.6	353	4	Q95144	Q95144 homo sapien
11	295.5	38.6	353	4	Q8NG07	Q8NG07 homo sapien
12	294.5	38.4	353	11	Q8VHZ4	Q8VHZ4 rattus norv
13	287	37.5	448	4	Q8NET3	Q8NET3 homo sapien
14	287	37.5	452	4	Q96PV4	Q96PV4 homo sapien
15	285.5	37.3	353	11	Q9CYP2	Q9CYP2 mus musculus
16	285.5	37.3	353	11	Q8C1C8	Q8C1C8 mus musculus

17	277	36.2	399	4	Q96A40	Q96A40 homo sapien
18	274	35.8	351	6	Q95KI4	Q95KI4 macaca fasc
19	271.5	35.4	351	4	Q96BY2	Q96BY2 homo sapien
20	266.5	34.8	237	4	Q9H833	Q9H833 homo sapien
21	266.5	34.8	351	4	Q9HAS1	Q9HAS1 homo sapien
22	257.5	33.6	352	11	Q9ERH6	Q9ERH6 mus musculus
23	250	32.6	402	11	Q9CZA5	Q9CZA5 mus musculus
24	247.5	32.3	403	4	Q8TE36	Q8TE36 homo sapien
25	245	32.0	393	11	Q9DB17	Q9DB17 mus musculus
26	245	32.0	393	11	Q8VD24	Q8VD24 mus musculus
27	237	30.9	402	4	Q8N1C1	Q8N1C1 homo sapien
28	100.5	13.1	386	4	Q8TCR7	Q8TCR7 homo sapien
29	98.5	12.9	525	11	Q80VR9	Q80VR9 mus musculus
30	97.5	12.7	218	4	Q9BT43	Q9BT43 homo sapien
31	97.5	12.7	225	11	Q8ROC0	Q8ROC0 mus musculus
32	97.5	12.7	543	4	Q8IU72	Q8IU72 homo sapien
33	92	12.0	185	16	Q8DE96	Q8DE96 vibrio vuln
34	92	12.0	286	16	Q8RG59	Q8RG59 fusbacteri
35	90.5	11.8	582	13	Q7ZU09	Q7ZU09 brachydanio
36	90	11.7	543	11	Q8KOZ0	Q8KOZ0 mus musculus
37	90	11.7	567	3	Q9UUS7	Q9UUS7 collatotic
38	90	11.7	592	4	Q9UJT2	Q9UJT2 homo sapien
39	89.5	11.7	218	11	Q9DBG0	Q9DBG0 mus musculus
40	89.5	11.7	471	10	Q93YN7	Q93YN7 arabidopsis
41	89.5	11.7	760	10	Q8VZU7	Q8VZU7 arabidopsis
42	89.5	11.7	800	4	O43290	O43290 homo sapien
43	89.5	11.7	1280	12	Q8QTF6	Q8QTF6 white spot
44	89	11.6	806	11	Q9Z314	Q9Z314 rattus norv
45	88.5	11.6	562	13	Q7SZP4	Q7SZP4 brachydanio

ALIGNMENTS

RESULT 1

Q9UL43 Q9UL43 PRELIMINARY; PRT; 149 AA.
AC Q9UL43, Q9UL43 (TREMREL. 13, Created)
DT 01-MAY-2000 (TREMREL. 13, Last sequence update)
DT 01-MAY-2000 (TREMREL. 13, Last annotation update)
DE Paraneoplastic cancer-testis-brain antigen (Fragment).
GN MA3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Gultekin S.H., Voltz R., Rosenfeld M.R., Gerstner E., Eichen J.,
RA Posner J.B., Dalmay J.;
RT "Identification of a novel cancer testis brain antigen using serum
antibodies from patients with testicular tumors and paraneoplastic
limbic encephalomyelitis";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF083114; AAF05625.1; -
FT NON TER
SQ SEQUENCE 149 AA; 17421 MW; D46136701D3C5651 CRC64;

Query Match 100.0%; Score 766; DB 4; Length 149;
Best Local Similarity 100.0%; Pred. No. 9.1e-62;
Matches 149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	DLHVIQADNPISVBECLFQVFGSLESRRTAQVRLKPYQEEGKVSAYVLRLETL	60
Db	1	DLHVIQADNPISVBECLFQVFGSLESRRTAQVRLKPYQEEGKVSAYVLRLETL	60
Qy	61	LRRAVEKRAIPRIADQVRLEQVMAGATLNQMLWCLRELKDGQPPPSFLELMKVIREEE	120
Db	61	LRRAVEKRAIPRIADQVRLEQVMAGATLNQMLWCLRELKDGQPPPSFLELMKVIREEE	120
Qy	121	EEEASFENESIEPERDGYGRWNHGGDD	149

```
Db 121 EEEASFENESIEPEERDGYGRWNHEGDD 149
RESULT 2
O94959 PRELIMINARY; PRT; 364 AA.
AC O94959;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein KIAA0883 (Paraneoplastic associated brain-testis-
DE cancer antigen).
GN KIAA0883.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=99156230; PubMed=10048485;
RA Nagase T., Ishikawa K., Suyama H., Kikuno R., Hirose M.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XII.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 5:335-364(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Eichen J.G., Dalmay J., Wade D., Rosenfeld M.R.;
RT "Characterization of a Brain-Testis-Cancer Antigen.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB020690; BAA74906.1; -.
DR EMBL; AF286487; AAG28165.1; -.
DR Genew; HGNC:9159; PNMA2.
DR InterPro; IPR005162; Retrotrans gag.
DR Pfam; PF03732; Retrotrans_gag; 1.
KW Hypothetical protein.
SQ SEQUENCE 364 AA; 41509 MW; 6E417AD96B3F0E93 CRC64;
Query Match 99.0%; Score 758; DB 4; Length 364;
Best Local Similarity 99.3%; Pred. No. 1.3e-60;
Matches 148; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 DLMHIVQADNPISVVEECLEAFKQVFGSLESRTAQVRYLKPQVEGKVSAYVLETL 60
Db 216 DLMHIVQADNPISVVEECLEAFKQVFGSLESRTAQVRYLKPQVEGKVSAYVLETL 275
QY 61 LRRAVEKRAIPRIADQVRLQVMAQATLNQMLWCLRELKDGQPPPSFLELMKVIREEE 120
Db 276 LRRAVEKRAIPRIADQVRLQVMAQATLNQMLWCLRELKDGQPPPSFLELMKVIREEE 335
QY 121 EEEASFENESIEPEERDGYGRWNHEGDD 149
Db 336 EEEASFENESIEPEERDGYGRWNHEGDD 364
RESULT 3
Q9UL42 PRELIMINARY; PRT; 283 AA.
AC Q9UL42;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Paraneoplastic cancer-testis-brain antigen (Fragment).
GN NA4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Gultekin S.H., Voltz R., Rosenfeld M.R., Gerstner E., Eichen J.,
```

```
RA Posner J.B., Dalmay J.;
RT "Identification of a novel cancer testis brain antigen using serum
RT antibodies from patients with testicular tumors and paraneoplastic
RT limbic encephalomyelitis.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF083115; AAF05626.1; -.
DR InterPro; IPR005162; Retrotrans gag.
DR Pfam; PF03732; Retrotrans_gag; 1.
FT NON_TER 1
SQ SEQUENCE 283 AA; 32333 MW; E27BA6BCDCD240A4 CRC64;
Query Match 98.6%; Score 755; DB 4; Length 283;
Best Local Similarity 98.7%; Pred. No. 1.9e-60;
Matches 147; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 DLMHIVQADNPISVVEECLEAFKQVFGSLESRTAQVRYLKPQVEGKVSAYVLETL 60
Db 135 DLMHIVQADNPISVVEECLEAFKQVFGSLESRTAQVRYLKPQVEGKVSAYVLETL 194
QY 61 LRRAVEKRAIPRIADQVRLQVMAQATLNQMLWCLRELKDGQPPPSFLELMKVIREEE 120
Db 195 LRRAVEKRAIPRIADQVRLQVMAQATLNQMLWCLRELKDGQPPPSFLELMKVIREEE 254
QY 121 EEEASFENESIEPEERDGYGRWNHEGDD 149
Db 255 EEEASFENESIEPEERDGYGRWNHEGDD 283
RESULT 4
Q9GMU3 PRELIMINARY; PRT; 364 AA.
AC Q9GMU3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
RT libraries.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047632; BAB12156.1; -.
DR InterPro; IPR005162; Retrotrans gag.
DR Pfam; PF03732; Retrotrans_gag; 1.
KW Hypothetical protein.
SQ SEQUENCE 364 AA; 41350 MW; 0CF72210D7EC1524 CRC64;
Query Match 96.7%; Score 741; DB 6; Length 364;
Best Local Similarity 97.3%; Pred. No. 4.7e-59;
Matches 145; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 DLMHIVQADNPISVVEECLEAFKQVFGSLESRTAQVRYLKPQVEGKVSAYVLETL 60
Db 216 DLMHIVQADNPISVVEECLEAFKQVFGSLESRTAQVRYLKPQVEGKVSAYVLETL 275
QY 61 LRRAVEKRAIPRIADQVRLQVMAQATLNQMLWCLRELKDGQPPPSFLELMKVIREEE 120
Db 276 LRRAVEKRAIPRIADQVRLQVMAQATLNQMLWCLRELKDGQPPPSFLELMKVIREEE 335
QY 121 EEEASFENESIEPEERDGYGRWNHEGDD 149
Db 336 EEEASFENESIEPEERDGYGRWNHEGDD 364
RESULT 5
```



```

Q8BHK0
ID Q8BHK0 PRELIMINARY; PRT; 365 AA.
AC Q8BHK0;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Similar to KIAA0883 protein.
DE PNA2 OR A830049P17RIK.
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cortex;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR ENBL; AK043718; BAC31626.1; -
DR ENBL; AK043910; BAC31700.1; -
DR MGD; MGI:2444129; Pnma2.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000425; MIP.
DR InterPro; IPR005162; Retrotrans_gag.
DR Pfam; PF03732; Retrotrans_gag; 1.
DR PROSITE; PS00221; MIP; 1.
SQ SEQUENCE 365 AA; 41201 MW; 0F90C940B9D843D9 CRC64;

Query Match 78.4%; Score 600.5; DB 11; Length 365;
Best Local Similarity 82.2%; Pred. No. 2.5e-46;
Matches 120; Conservative 11; Mismatches 14; Indels 1; Gaps 1;

QY 1 DLHIVQADNPISVVECLEAFKQVFGSLERRTAQVRYLKPQYEGEKVSAYVLRLETL 60
DB 216 DLHIVQADNPISVVGCELEAFKQVFGSTESRRTSQVKYLTQYQOGEKISAYVLRLETL 275

QY 61 LRAVEKRAIPRRIADQVRLEQVMAGATLNQMLWCLRELKDDQPPSPFLMKVIREEE 120
DB 276 LRAVEKRAIPRRIADQVRLEQVMAGANLGNVLWCLRELKDDQGLPTFLQLMKVIREEE 335

QY 121 EEE-ASPENESIEPERDGYGRWNH 145
DB 336 EEDAYPEQESREBPFGEREGSGCWN 361

RESULT 6
Q8JZW8
ID Q8JZW8 PRELIMINARY; PRT; 466 AA.
AC Q8JZW8;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Similar to paraneoplastic antigen MA3.
GN PNA3.
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RA Strausberg R.;
RL Submitted (AUG-2002) to the ENBL/GenBank/DBJ databases.
DR ENBL; BC036726; AAH36726.1; -
DR MGD; MGI:2180565; Pnma3.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008152; P:metabolism; IEA.

```

```

DR InterPro; IPR001544; AminoTrans_IV.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF00098; zf-CCHC; 1.
DR PRINTS; PR00939; C2HCZNFINGER.
DR PROSITE; PS0158; ZF_CCHC; 1.
SQ SEQUENCE 466 AA; 54040 MW; 113787E37B0B3AAD CRC64;

Query Match 42.2%; Score 323.5; DB 11; Length 466;
Best Local Similarity 49.6%; Pred. No. 4.3e-21;
Matches 64; Conservative 38; Mismatches 26; Indels 1; Gaps 1;

QY 2 LMHTVQADNPISVVECLEAFKQVFGSLERRTAQVRYLKPQYEGEKVSAYVLRLETL 61
DB 216 VVWVLRANNAITVKECLEALRVQFGSVDRKTIQLAFCKAYQEGKVSFVLRLETL 275

QY 62 RRAVEKRAIPRRIADQVRLEQVMAGATLNQMLWCLRELKDDQPPSPFLMKVIREEE 121
DB 276 QKALEKNAISKVNQVRLKRIILGGAILSAKLREKLKMLKQRRPPPGFLALVKLFRREEE 335

QY 122 E-EASFENE 129
DB 336 EWEATRGSE 344

RESULT 7
Q9H0A4
ID Q9H0A4 PRELIMINARY; PRT; 455 AA.
AC Q9H0A4;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Hypothetical protein.
GN DKFZP434K225.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=21154917; PubMed=11230166;
RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
RA Anorge W., Boecker M., Bloecker H., Bauersachs S., Blum H.,
RA Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,
RA Mewes H.W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,
RA Wambutt R., Korn B., Klein M., Poustka A.;
RT "Towards a Catalog of Human Genes and Proteins: Sequencing and
RT Analysis of 500 Novel Complete Protein Coding Human cDNAs.";
RL Genome Res. 11:422-435(2001).
DR ENBL; AL136878; CAB66812.1; -
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF00098; zf-CCHC; 1.
DR PRINTS; PR00939; C2HCZNFINGER.
DR SMART; SM00343; ZNF_C2HC; 1.
DR PROSITE; PS0158; ZF_CCHC; 1.
KW Hypothetical protein.
SQ SEQUENCE 455 AA; 51514 MW; 9AC8CC06685A8956 CRC64;

Query Match 42.2%; Score 323; DB 4; Length 455;
Best Local Similarity 52.7%; Pred. No. 4.6e-21;
Matches 68; Conservative 26; Mismatches 33; Indels 2; Gaps 1;

QY 6 VQADNPISVVECLEAFKQVFGSLERRTAQVRYLKPQYEGEKVSAYVLRLETLRAV 65
DB 220 LRASNASITVEECLEALQVFGVPVESHKIAQVKLCKAYQEGEKVSFVLRLEPLQRAV 279

QY 66 EKRAIPRRIADQVRLEQVMAGATLNQMLWCLRELKDDQPPSPFLMKVIREEESEAS 125
DB 280 ENNVYSRRNVNQTRLKRVLSGATLPDKLRDLKLMKQRRPPPGFLALVKLREEEWEAT 339

QY 126 F--ENESIE 132

```

```

Db          340 LGPDRESLE 348

RESULT 8
Q9UL41
ID Q9UL41 PRELIMINARY; PRT; 463 AA.
AC Q9UL41;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Paraneoplastic neuronal antigen MA3.
GN MA3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Gultekin S.H., Voltz R., Rosenfeld M.R., Gerstner E., Eichen J.,
RA Posner J.B., Dalmau J.;
RT "Identification of a novel cancer testis brain antigen using serum
RT antibodies from patients with testicular tumors and paraneoplastic
RT limbic encephalomyelitis.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF083116; AAF05627.1; -.
DR Genew; HGNC:18742; PNMA3.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF00098; zf_CCHC; 1.
DR PRINTS; PR00939; C2HCZNFINGER.
DR SMART; SM00343; Znf_C2HC; 1.
DR PROSITE; PS50158; ZF_CCHC; 1.
DR SEQUENCE 463 AA; 52376 MW; 0843EF6601D8557C CRC64;

Query Match 42.2%; Score 323; DB 4; Length 463;
Best Local Similarity 52.7%; Pred. No. 4.7e-21;
Matches 68; Conservative 26; Mismatches 33; Indels 2; Gaps 1;

QY 6 VQADNPISVEECLEAFKQVFGSLESRRTAQVRLKPYQEEGKVSAYVLRLETLRAV 65
Db 220 LRASNASITVEECLEAALQVFGVESHKIAQVKLCRAQYQEGKVSFVLRLEPLQRAV 279

QY 66 EKRAIPRRADQVRLQVFNAGATLNQMLWCLRELKDGPPSPFLELMKVIREEEEAS 125
Db 280 ENNVSRNNVNTLKKVLSGATLPDKLRDKLKMQRKPPGFLLVLRLEEEWEAT 339

QY 126 F--ENESIE 132
Db 340 LGPDRESLE 348

RESULT 9
Q8ND90
ID Q8ND90 PRELIMINARY; PRT; 192 AA.
AC Q8ND90;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein (Fragment).
GN DKFZP434N171.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Poustka A., Wellenreuther R., Mewes H.W., Weil B., Wiemann S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL834327; CAD38995.1; -.
KW Hypothetical protein.
FT NON_TER 1
FT SEQUENCE 192 AA; 21456 MW; 5DB17CF6A760A628 CRC64;

Query Match 38.6%; Score 295.5; DB 4; Length 192;
Best Local Similarity 46.3%; Pred. No. 5.4e-19;
Matches 62; Conservative 29; Mismatches 36; Indels 7; Gaps 2;

QY 1 DLMHIVQADNPISVEECLEAFKQVFGSLESRRTAQVRLKPYQEEGKVSAYVLRLETL 60
Db 212 DVIRILKSNPAITTAECLEKALEQVFGSVESRRDAQIKFLNTYQNFGEKLSAVIRLEPL 271

QY 61 LRRAVEKRAIPRIADQVRLQVFMAGA---TLNQMLWCLRELKDGPPSPFLELMKVI 116
Db 272 LQKVVEKGAIKDNVNQARLEQVIAGANHSAIRRLQWL---TGAGEGPAPNLFQLLVQI 328

QY 117 REEEEEEASFENES 130
Db 329 REEEAKEEEEEAEA 342

RESULT 11
Q8NG07

```


Query Match 37.5%; Score 287; DB 4; Length 452;
Best Local Similarity 43.8%; Pred. No. 8.4e-18;
Matches 63; Conservative 30; Mismatches 43; Indels 8; Gaps 2;

QY 2 LMHVQADNPISVBECEAFKQVFGSLESRTAQVRLKPYQEGEKVSAYVLRLETL 61
DB 215 IMRVLAQNNDSITVQCCLDAKQIFGKEDPRASQFRFLQTSFKIGEKVSTELLRLPL 274
QY 62 RRAVEKRAIPRIADQVRLQVWAGATLNQMLWCRRLKLDQGGPPSFLMLKVIREEE 121
DB 275 QKAVHKSPLSVRSTDMIRKLHLLARVAMTPALRGKLELLDQRCPPNFLELMKLRDEEE 334
QY 122 ---EASPENESIEPPEERDGYGR 142
DB 335 WENTAVMKNK-----EKPSGRGR 353

Search completed: September 27, 2004, 17:10:09
Job time : 20.7467 secs

RESULT 15
Q9CYP2 PRELIMINARY; PRT; 353 AA.

AC Q9CYP2;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE 5730402C15RIK protein.
GN PNMA1 OR 5730402C15RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischer W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayaishizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL; AK017476; BAB30762.1; --
DR MGD; MGI:2180564; Pnmal.
SQ SEQUENCE 353 AA; 39718 MW; F7B27378B8469675 CRC64;

Query Match 37.3%; Score 285.5; DB 11; Length 353;
Best Local Similarity 47.7%; Pred. No. 8.7e-18;
Matches 63; Conservative 25; Mismatches 33; Indels 11; Gaps 3;

QY 1 DLMHVQADNPISVBECEAFKQVFGSLESRTAQVRLKPYQEGEKVSAYVLRLETL 60
DB 212 DVIRLTKNTAAITAECKALQVFGSVSSSDAQVRFLNTYQNGEKLSSVIRLEPL 271
QY 61 LRRAVEKRAIPRIADQVRLQVWAGATLNQMLWCRRLKLDQGGPPSFLMLKVI 116
DB 272 LQKVDKGVIDKDNVQARLEQVIAGANHSGLRRQLWLAGAE---EGPAPNLFQLLVQI 328
QY 117 REE-----EEEA 124
DB 329 REEAKKEEEEA 340

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 27, 2004, 17:04:25 ; Search time 4.86928 Seconds
(without alignments)
1593.347 Million cell updates/sec

Title: US-10-037-860-9

Perfect score: 766

Sequence: 1 DLMHIVQADNPSISVEECLE.....SIEEPERDGYGRWNHEGDD 149

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	758	99.0	364	1	PMA2_HUMAN	Q9ul42 homo sapien
2	741	96.7	364	1	PMA2_MACFA	Q9gm33 macaca fasc
3	600.5	78.4	365	1	PMA2_MOUSE	Q8bhk0 mus musculus
4	295.5	38.6	353	1	PMA1_HUMAN	Q8nd90 homo sapien
5	294.5	38.4	353	1	PMA1_RAT	Q8vth4 rattus norv
6	285.5	37.3	353	1	PMA1_MOUSE	Q8cic8 mus musculus
7	274	35.8	351	1	MOPI_MACFA	Q95ki4 macaca fasc
8	271.5	35.4	351	1	MOPI_HUMAN	Q96by2 homo sapien
9	257.5	33.6	352	1	MOPI_MOUSE	Q9erh6 mus musculus
10	102.5	13.4	554	1	NFL_BOVIN	P02548 bos taurus
11	98.5	12.9	525	1	NAB2_MOUSE	Q61127 mus musculus
12	96.5	12.6	541	1	NFL_RAT	P19527 rattus norv
13	96	12.5	544	1	NFL_XENLA	P35616 xenopus lae
14	95	12.4	759	1	CHLD_ARATH	Q98je1 arabidopsis
15	94.5	12.3	525	1	NAB2_HUMAN	Q15742 homo sapien
16	94	12.3	548	1	NFL_PIG	P02547 sus scrofa
17	94	12.3	1330	1	KTNI_VULVU	Q97961 vulpes vulp
18	93.5	12.2	543	1	NFL_HUMAN	P07196 homo sapien
19	93.5	12.2	1364	1	CND1_XENLA	Q9yhy6 xenopus lae
20	93	12.1	399	1	RPSD_THEMA	P77994 thermotoga
21	91	11.9	657	1	BFS1_CHICK	Q06637 gallus gall
22	90.5	11.8	582	1	PESC_BRARE	P79741 brachydanio
23	90	11.7	542	1	NFL_MOUSE	P08551 mus musculus
24	85	11.1	387	1	NUS4_THETH	P48514 thermus the
25	83.5	10.9	631	1	RA21_HUMAN	P60216 homo sapien
26	83.5	10.9	776	1	TOPI_RICPR	Q9zdk2 rickettsia
27	83	10.8	1392	1	CND1_MOUSE	Q8k224 mus musculus
28	83	10.8	1962	1	MYS4_DROME	P05661 drosophila
29	82	10.7	1225	1	SA3_HUMAN	Q9uj98 homo sapien
30	82	10.7	1357	1	KTNI_HUMAN	Q86up2 homo sapien
31	81.5	10.6	105	1	RL31_AERPE	Q9ydz5 aeropyrum p
32	81.5	10.6	1898	1	TRHY_HUMAN	Q07283 homo sapien
33	80.5	10.5	1217	1	TOPI_ZYMMO	Q9x3x7 zymomonas m

RESULT 1									
PMA2_HUMAN									
ID	PMA2_HUMAN	STANDARD;	PRT;	364	AA.				
AC	Q9UL42; Q94959; Q95145; Q9UL43;								P46060 homo sapien
DT	15-MAR-2004 (Rel. 43, Created)								Q9usi6 schizosacch
DT	15-MAR-2004 (Rel. 43, Last sequence update)								Q9hh12 thermococcu
DT	15-MAR-2004 (Rel. 43, Last annotation update)								P74557 synechocyst
DE	Paraneoplastic antigen M2 (Onconeural antigen MA2) (Paraneoplastic neuronal antigen MM2) (40 kDa neuronal protein).								P37963 bacillus su
GN	PMA2 OR MA2 OR KIA0883.								Q58914 methanococc
OS	Homo sapiens (Human).								P53252 saccharomyc
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;								Q9byv6 homo sapien
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.								O60841 homo sapien
OX	NCBI_TaxID=9606;								Q9gy06 mus musculu
RN	[1]								Q7nvq6 porphyromon
RP	SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.								Q02916 coturnix co
RC	TISSUE=Cerebellum;								
RX	MEDLINE=9270611; PubMed=10362822;								
RA	Voltz R., Gultekin S.H., Rosenfeld M.R., Gerstner E., Eichen J.,								
RA	Posner J.B., Daimau J.;								
RT	"A serologic marker of paraneoplastic limbic and brain-stem								
RT	encephalitis in patients with testicular cancer.";								
RL	New Engl. J. Med. 340:1788-1795(1999).								
RN	[2]								
RP	SEQUENCE FROM N.A.								
RC	TISSUE=Brain;								
RX	MEDLINE=99156230; PubMed=10048485;								
RA	Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirotsawa M.,								
RA	Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;								
RT	"Prediction of the coding sequences of unidentified human genes. XII.								
RT	The complete sequences of 100 new cDNA clones from brain which code								
RL	for large proteins in vitro.";								
RN	DNA Res. 5:355-364(1998).								
RN	[3]								
RP	IDENTIFICATION, AND SUBCELLULAR LOCATION.								
RP	MEDLINE=99158179; PubMed=10050892;								
RA	Dalmay J., Gultekin S.H., Voltz R., Hoard R., DesChamps T.,								
RA	Balmaceda C., Batchelor T., Gerstner E., Eichen J., Frennieri J.,								
RA	Posner J.B., Rosenfeld M.R.;								
RT	"Mal, a novel neuron- and testis-specific protein, is recognized by								
RT	the serum of patients with paraneoplastic neurological disorders.";								
RL	Brain 122:27-39(1999).								
CC	-1- SUBCELLULAR LOCATION: Nuclear; nucleolar.								
CC	-1- TISSUE SPECIFICITY: Brain specific. In some patients suffering								
CC	from cancers, it is also specifically expressed by the testicular								
CC	tumor cells.								
CC	-1- MISCELLANEOUS: Antibodies against PMA2 are present in sera from								
CC	patients suffering of paraneoplastic neurological disorders.								
CC	-1- SIMILARITY: Belongs to the PNMA family.								
CC	-----								
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration								
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -								
CC	the European Bioinformatics Institute. There are no restrictions on its								
CC	use by non-profit institutions as long as its content is in no way								
CC	modified and this statement is not removed. Usage by and for commercial								
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/								

or send an email to license@isb-sib.ch.

```

CC EMBL; AF037365; AAD02098.1; -
CC EMBL; AF083114; AAF05625.1; -
DR EMBL; AF083115; AAF05626.1; -
DR EMBL; AF286487; AAG28165.1; -
DR EMBL; AB020690; BAA74906.1; -
DR EMBL; HGNC:9159; PNM2.
DR MIM; 603970; -
KW Antigen; Tumor antigen; Nuclear protein.
FT DOMAIN 333 338 POLY-GLU.
FT CONFLICT 128 129 GV -> AL (IN REF. 1; AAD02098).
FT CONFLICT 141 141 L -> I (IN REF. 1; AAD02098).
FT CONFLICT 257 257 T -> P (IN REF. 1; AAF05625).
FT CONFLICT 278 278 R -> K (IN REF. 1; AAF05626).
SQ SEQUENCE 364 AA; 41509 MW; 6E417AD96E3F0E93 CRC64;

Query Match
Best Local Similarity 99.0%; Score 758; DB 1; Length 364;
Matches 148; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DLMHIVQADNPSISVEECLEAFKQVFGSLESRRTAQVRYLKPQYQEGEKVSAYVLRLETL 60
DB 216 DLMHIVQADNPSISVEECLEAFKQVFGSLESRRTAQVRYLKPQYQEGEKVSAYVLRLETL 275

QY 61 LRRAVEKRAIPRIADQVRLQVMAQATLNQMLWCLRELKDKQGGPPSPFLELMKVIREEE 120
DB 276 LRRAVEKRAIPRIADQVRLQVMAQATLNQMLWCLRELKDKQGGPPSPFLELMKVIREEE 335

QY 121 EEEASFENESIEPEERDGYGRWNHEGDD 149
DB 336 EEEASFENESIEPEERDGYGRWNHEGDD 364

RESULT 2
PMA2_MOUSE STANDARD; PRT; 365 AA.
ID PMA2_MOUSE
AC Q8BHK0:
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Paraneoplastic antigen Ma2 homolog.
GN PMA2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Brain cortex;
RX MEDLINE=22354683; PubMed=12466851;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konoaga A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Mahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang N.,
RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda I.,
RA Hara A., Hashizume W., Imctani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Sasaki D., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RA "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
CC -1- SUBCELLULAR LOCATION: Nuclear; nucleolar (By similarity).
CC -1- SIMILARITY: Belongs to the PMA family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; AB047632; BAB12156.1; -
KW Nuclear protein.
FT DOMAIN 333 POLY-GLU.
FT SEQUENCE 364 AA; 41350 MW; 0CF72210D7EC1524 CRC64;

Query Match
Score 741; DB 1; Length 364;

```

```

DR EMBL; AK043718; BAC31626.1; -
DR EMBL; AK043910; BAC31700.1; -
DR MGD; MGI:2444129; Pnma2.
KW Nuclear protein.
FT DOMAIN 333 338 POLY-GLU.
SQ SEQUENCE 365 AA; 0F90C940B9D843D9 CRC64;

Query Match 78.4%; Score 600.5; DB 1; Length 365;
Best Local Similarity 82.2%; Pred. No. 1.1e-42;
Matches 120; Conservative 11; Mismatches 14; Indels 1; Gaps 1;

Qy 1 DLHMIVQADNPISVVEECLEAFKQVFGSLERRTAQVRYLKPYQEGEKVSAYVLRLETL 60
Db 216 DLHMIVQADNPISVSGECLEAFKQVFGSTERRISQVILYTKYQEGEKISAYVLRLETL 275
Qy 61 LRRAVEKRAIPRRIADQVRLQVQVAGATINQMLRCRLELKDQGPSPFLELMKVIREE 120
Db 276 LRRAVEKRAIPRRIADQVRLQVQVAGATINQMLRCRLELKDQGPSPFLELMKVIREE 335
Qy 121 EEE-ASPENSIEPEERDGYGRNH 145
Db 336 EEDDAYFEQSRREPGEREGSGCWN 361

RESULT 4
PMAL_HUMAN
ID PMAL_HUMAN STANDARD; PRT; 353 AA.
AC Q8ND90; O95144; Q8NG07;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Paraneoplastic antigen Mal (Neuron- and testis-specific protein 1)
DE (37 kDa neuronal protein).
DE PNMA1 OR MAL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Carnathini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RC TISSUE=Cerebellum;
RX MEDLINE=99158179; PubMed=10050892;
RA Balmau J., Gultekin S.H., Voltz R., Hoard R., DesChamps T.,
RA Balmaceda C., Batchelor T., Gersner E., Eichen J., Fremier J.,
RA Posner J.B., Rosenfeld M.R.;
RT "Mal, a novel neuron- and testis-specific protein, is recognized by
RT the serum of patients with paraneoplastic neurological disorders.";
RL Brain 122:27-39(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Schueller M.M., Jenne D.E., Schutze dit Belkner N., Hohlfield R.,
RA Voltz R.;
RT "Mal (PNMA1).";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Narusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.D., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Rulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

```

```

RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP SEQUENCE OF 162-353 FROM N.A.
RC TISSUE=Testis;
RA Poustka A., Wellenreuther R., Mewes H.-W., Weil B., Wiemann S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Nuclear; nucleolar. In tumor cells, it is
CC cytoplasmic.
CC -1- TISSUE SPECIFICITY: Testis and brain specific. In some patients
CC suffering from cancers, it is also specifically expressed by the
CC paraneoplastic tumor cells.
CC -1- MISCELLANEOUS: Antibodies against PNMA1 are present in sera from
CC patients suffering of paraneoplastic neurological disorders.
CC -1- SIMILARITY: Belongs to the PNMA family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF037364; AAD13810.3; -
DR EMBL; AF320308; AAN05100.1; -
DR EMBL; BC039577; AAH39577.1; -
DR EMBL; AL834327; CAD38995.1; -
DR Genew; HGNC:9158; PNMA1.
DR MIM; 604010; -
DR GO; GO:0005737; C:cytoplasm; TAS.
DR GO; GO:0005730; C:nucleolus; TAS.
DR GO; GO:0007417; P:central nervous system development; TAS.
DR GO; GO:0007283; P:spermatogenesis; TAS.
KW Antigen; Tumor antigen; Nuclear protein.
FT DOMAIN 335 341 POLY-GLU.
SQ SEQUENCE 353 AA; 39761 MW; EB7F5B6AEDA25961 CRC64;

Query Match 38.6%; Score 295.5; DB 1; Length 353;
Best Local Similarity 46.3%; Pred. No. 1.9e-17;
Matches 62; Conservative 29; Mismatches 36; Indels 7; Gaps 2;

Qy 1 DLHMIVQADNPISVVEECLEAFKQVFGSLERRTAQVRYLKPYQEGEKVSAYVLRLETL 60
Db 212 DVIRILKSNPNAITTAECLEKALEQVFGSVESRDQIKFLNTYQNPGEKLSAYVLRLEPL 271
Qy 61 LRRAVEKRAIPRRIADQVRLQVQVAGATINQMLRCRLELKDQGPSPFLELMKV 116
Db 272 LQKVVEKGAIDKQNVNOARLEQVIAGNHSGAIRQLWL---TGAGEGPAPNLFQLLVQI 328
Qy 117 REEEEEEASFENES 130
Db 329 REEEAKSEEEEA 342

RESULT 5
PMAL_RAT
ID PMAL_RAT STANDARD; PRT; 353 AA.
AC Q8VHZ4;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Paraneoplastic antigen Mal homolog.
GN PNMA1 OR MAL.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.

```



```

RX MEDLINE=99158179; PubMed=10050892;
RA Dalmau J., Gultekin S.H., Voltz R., Hoard R., DesChamps T.,
RA Balmaceda C., Batchelor T., Gerstner E., Eichen J., Frennier J.,
RA Posner J.B., Rosenfeld M.R.;
RT "Maf, a novel neuron- and testis-specific protein, is recognized by
RT the serum of patients with paraneoplastic neurological disorders.";
RL Brain 122:27-39(1999).
CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar (By similarity).
CC -!- TISSUE SPECIFICITY: Testis and brain specific.
CC -!- SIMILARITY: Belongs to the PNA family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF335505; AAL73196.1; -.
KW Nuclear protein.
FT DOMAIN 335 341 POLY-GLU.
SQ SEQUENCE 353 AA; 39793 MW; PFC3A717FDCEDA33 CRC64;

Query Match 38.4%; Score 294.5; DB 1; Length 353;
Best Local Similarity 47.0%; Pred. No. 2.3e-17;
Matches 63; Conservative 27; Mismatches 37; Indels 7; Gaps 2;

QY 1 DLMHVQADNPSTSVBECLEAFKQVFGSLESRTAQVRLKPYQEGEKVSAYVLRLETL 60
Db 117 REE-----EEEE 124
QY 117 REE-----EEEE 124
Db 329 REEAAKEEEEA 342

RESULT 6
PMAL_MOUSE
ID PMAL_MOUSE STANDARD; PRT; 353 AA.
AC Q8C1C8; Q9CVP2;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Paraneoplastic antigen Maf homolog.
GN PMAL OR MAF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Brain, and Embryo;
RX MEDLINE=22354683; PubMed=12466851;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Balderelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.L.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Nagtogh D.R., Numaiz L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Nakata K., Okido T., Pavan W.J., Fertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,

```

```

RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shingawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar (By similarity).
CC -!- SIMILARITY: Belongs to the PNA family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AK017476; BAB30762.1; -.
DR EMBL; AK028331; BAC25885.1; -.
DR MGD; MGI:2180564; Pnmal.
KW Nuclear protein.
FT DOMAIN 336 341 POLY-GLU.
FT CONFLICT 136 136 T -> M (IN REF. 1; BAB30762).
SQ SEQUENCE 353 AA; 39688 MW; 18CEDC3AC4E70939 CRC64;

Query Match 37.3%; Score 285.5; DB 1; Length 353;
Best Local Similarity 47.7%; Pred. No. 1.3e-16;
Matches 63; Conservative 25; Mismatches 33; Indels 11; Gaps 3;

QY 1 DLMHVQADNPSTSVBECLEAFKQVFGSLESRTAQVRLKPYQEGEKVSAYVLRLETL 60
Db 212 DVIRILKTNNAITTAECLEQVFGSVESRDQVRFLNTYQNPGEKLSVVIRLEPL 271
QY 61 LRRAVEKRAIPRIADQVRLQVMAQA-----TINQMLWCRRLKDKQGGPPPSFLELMKVI 116
Db 272 LQKVVDKGVIDKDNVQARLEQVIAGANHSGLRRQLWLAGAE---EGPAPNLFQLLVQI 328
QY 117 REE-----EEEE 124
Db 329 REEAAKEEEEA 340

RESULT 7
MOP1_MACFA
ID MOP1_MACFA STANDARD; PRT; 351 AA.
AC Q95KI4;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Modulator of apoptosis 1 (MAP-1).
GN MOAP1.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Temporal cortex;
RX MEDLINE=12354683; PubMed=12466851;
RA Oseada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
RT libraries.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Mediates caspase-dependent apoptosis (By similarity).

```



```

QY 118 E---EEEEE 124
DB 331 DYEAEEEEE 340

RESULT 9
MOPI MOUSE STANDARD; PRT; 352 AA.
AC Q9EH6;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DE 15-MAR-2004 (Rel. 43, Last annotation update)
DE Modulator of apoptosis 1 (MAP-1).
GN MOAP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21264738; PubMed=11060313;
TX Tan K.O., Tan K.M.L., Chan S.-L., Yee K.S.Y., Bevort M., Ang K.C.,
TX Yu V.C.;
RT "MAP-1, a novel proapoptotic protein containing a BH3-like motif that
RT associates with Bax through its Bcl-2 homology domains.";
RL J. Biol. Chem. 276:2802-2807(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.P.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez J., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6NCr; TISSUE=Hematopoietic, and Kidney;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altachul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.L., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.J., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human

```

```

RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -|- FUNCTION: Mediates caspase-dependent apoptosis.
CC -|- SUBUNIT: Homodimer. Binds BAX, BCL2 and BCLX (By similarity).
CC -|- DOMAIN: The BH3-like domain is required for association with BAX
CC and for mediating apoptosis. The three BH domains (BH1, BH2, and
CC BH3) of BAX are all required for mediating protein-protein
CC interaction (By similarity).
CC -|- SIMILARITY: Belongs to the PNMA family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF305551; AAG31787.1; -.
CC DR EMBL; AK019599; BAB31810.1; -.
CC DR EMBL; BC014715; AAH14715.1; -.
CC DR EMBL; BC055374; AAH55374.1; -.
CC DR MGI; MGI:1915555; Moapl.
KW Apoptosis.
FT SITE 120 127 BH3-LIKE.
FT DOMAIN 335 340 POLY-GLU.
FT CONFLICT 57 57 R -> K (IN REF. 3; AAH55374).
SQ SEQUENCE 352 AA; 39404 MW; 8F4630D080495D98 CRC64;

Query Match 33.6%; Score 257.5; DB 1; Length 352;
Best Local Similarity 38.6%; Pred. No. 2.7e-14;
Matches 54; Conservative 35; Mismatches 44; Indels 7; Gaps 2;

QY 1 DLMHIVQADNPISVVEECLEAKQVFGSLESRTAQVRLKPYQEGSEKVSAYVLRLETL 60
DB 218 EIIRVLKINNPFITVAECUKLTETIFGIIDNRAIQVRYLTYYQKTEKLSAYVLRLEPL 277
QY 61 LRRRAVEKRAIPRIADQVRLQVMAQATLQMLWCLRELKQDQPPPSFLELMKVIREEE 120
DB 278 LQKLQVKGAIKEVNVQARLDQVIAGAVHKSVR--RELGLPEGSPAPGLQLLTLYDKKE 335
QY 121 EEEASFENESIEPEERDGY 140
DB 336 AEE-----EVLVLAELGY 350

RESULT 10
NFL BOVIN STANDARD; PRT; 554 AA.
AC P02548; P79127;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neurofilament triplet L protein (68 kDa neurofilament protein)
DE (Neurofilament light polypeptide) (NF-L) (Micro glutamic acid-rich
DE protein).
GN NEFL.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Holstein; TISSUE=Brain;
RA Hill W.D., Zhang L., Balin B.J., Sprinkle T.J.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 468-554.
RX MEDLINE=85154567; PubMed=3884373;
RA Isobe T., Okuyama T.;
RT "Brain micro glutamic acid-rich protein is the C-terminal endpiece of
RT the neurofilament 68-kDa protein as determined by the primary

```

```

RT sequence.";
RL FEBS Lett. 182:389-392(1985).
CC -!- FUNCTION: Neurofilaments usually contain three intermediate
CC filament proteins: L, M, and H which are involved in the
CC maintenance of neuronal caliber.
CC -!- DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH
CC THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT
CC PROTEINS ARE DUE TO THE TAILPIECE EXTENSIONS. THIS REGION MAY FORM
CC A CHARGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH
CC OTHER NEURONAL COMPONENTS OR IONS.
CC -!- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE
CC NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL
CC INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM
CC FILAMENTS.
CC -!- SIMILARITY: Belongs to the intermediate filament family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U83919; AAB41543.1; -.
CC InterPro; IPR006821; Filament_head.
CC InterPro; IPR001664; IF.
CC Pfam; PF00038; filament; 1.
CC Pfam; PF04732; filament_head; 1.
CC PROSITE; PS00226; IF; 1.
CC Intermediate filament; Coiled coil; Neutrone.
CC INIT_MET 0 0
CC DOMAIN 1 92
CC DOMAIN 93 396
CC DOMAIN 397 554
CC DOMAIN 554 124
CC DOMAIN 125 137
CC DOMAIN 138 233
CC DOMAIN 234 252
CC DOMAIN 253 271
CC DOMAIN 272 280
CC DOMAIN 281 396
CC DOMAIN 397 443
CC DOMAIN 444 554
CC TAIL, SUBDOMAIN A.
CC TAIL, SUBDOMAIN B. (ACIDIC).
CC MISSING (IN REF. 2).
CC A -> AEA (IN REF. 2).
CC SEQUENCE 554 AA; 62514 MW; D772B81CA2C31CIA CRC64;
Query Match 13.4%; Score 102.5; DB 1; Length 554;
Best Local Similarity 23.0%; Pred. No. 0.31;
Matches 41; Conservative 24; Mismatches 66; Indels 47; Gaps 6;
Qy 2 LMHIVQADNPISIVVECLAFKQVFGSLRSRTAQVRLKPYQE----- 45
Db 333 LQELDKQADISAMQ--DTINKLENLRTTKSEMARYLKQYDQLLNVMKALDIEIAAYR 390
Qy 46 ---RGEX-----VSAVVLRLLETLRRA-----VEKRAIPRIADQVRL 81
Db 391 KLLGEETRLSFTSGVSLTTQTOSQVFGSAYGGLQTSYLSMSARSPFPYTSYHQVE 450
Qy 82 QVMAGATLNQMLCRRLKQDQPPGPFLELMKVIREEESEAESEFENESIEPEPERDG 139
Db 451 QIEVEETIEA---AKAEAEKDE--PPSEGEAEKEKEAEAEAEAEAEAEAE 503
RESULT 11
ID_NAB2_MOUSE STANDARD; PRT; 525 AA.
AC Q61127;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE NGFI-A binding protein 2 (EGR-1 binding protein 2).

```

```

GN NAB2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Brain;
RX MEDLINE=96251303; PubMed=8668170;
RA Svaren J., Severson B.R., Apel E.D., Zimonjic D.B., Popescu N.C.,
RA Milbrandt J.;
RT "NAB2, a corepressor of NGFI-A (Egr-1) and Krox20, is induced by
RT proliferative and differentiative stimuli.";
RL Mol. Cell. Biol. 16:3545-3553(1996).
RN [2]
RP PARTIAL SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=97271553; PubMed=9126479;
RA Svaren J., Apel E.D., Simburger K.S., Jenkins N.A., Gilbert D.J.,
RA Copeland N.A., Milbrandt J.;
RT "The Nab2 and Stat6 genes share a common transcription termination
RT region.";
RL Genomics 41:33-39(1997).
CC -!- FUNCTION: Acts as a transcriptional repressor for zinc finger
CC transcription factors EGR1 and EGR2. Isoform 2 lacks repression
CC ability. Homomultimeric may associate with EGR1 bound to DNA (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: NUCLEAR. ISOFORM 2 IS NOT LOCALIZED TO THE
CC NUCLEUS.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q61127-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q61127-2; Sequence=VSP_003388, VSP_003389;
CC -!- TISSUE SPECIFICITY: Highly expressed in brain and thymus, and at
CC lower levels in spleen, kidney, heart and testis. Isoform 1 is
CC predominantly expressed in testis, whereas isoform 3 is more
CC abundant in thymus.
CC -!- INDUCTION: By serum stimulation.
CC -!- DOMAIN: The NAB conserved domain 1 (NCD1) interacts with EGR1
CC inhibitory domain and mediates multimerization.
CC -!- DOMAIN: The NAB conserved domain 2 (NCD2) is necessary for
CC transcriptional repression.
CC -!- SIMILARITY: BELONGS TO THE NAB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U47543; AAC52650.1; -.
CC MGI; MGI:107563; Nab2.
CC InterPro; IPR006989; Nab_cent.
CC InterPro; IPR006988; Nab_N.
CC Pfam; PF04904; NCD1; 1.
CC Pfam; PF04905; NCD2; 1.
CC Transcription regulation; Repressor; Alternative splicing.
CC DOMAIN 35 113 NCD1.
CC DOMAIN 267 356 NCD2.
CC DOMAIN 353 384
CC VARSPLIC 320 322 LTI -> ASL (in isoform 2).
CC VARSPLIC 323 525 /FTId=VSP_003388.
CC VARSPLIC 323 525 Missing (in isoform 2).
CC VARSPLIC 323 525 /FTId=VSP_003389.
CC SEQUENCE 525 AA; 56653 MW; FED428E9A8BD804 CRC64;
Query Match 12.9%; Score 98.5; DB 1; Length 525;
Best Local Similarity 24.2%; Pred. No. 0.62;

```


10-OCT-2003 (Rel. 42, Last annotation update)
 Neurofilament triplet L protein (Neurofilament light polypeptide) (NP-L).
 Xenopus laevis (African clawed frog).
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae; Xenopus.
 NCBI_TaxID=8355;
 RN [1]
 SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=92356194; PubMed=1494944;
 RA Charnas L.R., Szaro B.G., Gainer H.;
 RT "Identification and developmental expression of a novel low molecular weight neuronal intermediate filament protein expressed in Xenopus laevis.";
 RL J. Neurosci. 12:3010-3024(1992).
 CC -!- FUNCTION: Neurofilaments usually contain three intermediate filament proteins: L, M, and H which are involved in the maintenance of neuronal caliber.
 CC -!- MISCELLANEOUS: NP-L IS THE MOST ABUNDANT OF THE THREE NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM FILAMENTS.
 CC -!- SIMILARITY: Belongs to the intermediate filament family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL; M86654; AA83018.1; -;
 DR InterPro; IPR006821; Filament_head.
 DR InterPro; IPR001664; IF.
 DR Pfam; PF00038; filament; 1.
 DR Pfam; PF04732; filament_head; 1.
 DR PROSITE; PS00226; IF; 1.
 KW Intermediate filament; Coiled coil; Neurone.
 FT DOMAIN 1 87 HEAD.
 FT DOMAIN 88 390 ROD.
 FT DOMAIN 391 544 TAIL.
 FT DOMAIN 88 119 COIL 1A.
 FT DOMAIN 120 132 COIL 1B.
 FT DOMAIN 133 228 COIL 1C.
 FT DOMAIN 229 246 COIL 2A.
 FT DOMAIN 247 265 COIL 2B.
 FT DOMAIN 266 274 COIL 2C.
 FT DOMAIN 275 390 COIL 2D.
 FT DOMAIN 391 435 TAIL, SUBDOMAIN A.
 FT DOMAIN 436 544 TAIL, SUBDOMAIN B (ACIDIC).
 FT DOMAIN 441 538 GLU-RICH.
 FT DOMAIN 464 469 POLY-GLU.
 SQ SEQUENCE 544 AA; 61861 MW; 76D911B896E97201 CRC64;
 Query Match 12.5%; Score 96; DB 1; Length 544;
 Best Local Similarity 23.0%; Pred. No. 1;
 Matches 42; Conservative 19; Mismatches 56; Indels 66; Gaps 6;
 9 DNPISVEECLEAFQVFGSLESRTAQVRLKPYQE-----RGEK 49
 332 DKQGEIAGQMDINKLEELRLNTSEMYRLKQYDQLLNKVMALDIEIAYRKLLEGEE 391
 50 V-----SAYVRLTLLRRAVEKRAIPRIADQVRLQVWAG 86
 392 TRLSFGVGAIATGYTQAPVFGSAYSLQSSVM-----TSRAPTYTSHVQEEQDIE 447
 87 ATLNMQLWCLRLKLDGQPPSPFLMKVIRREBEFEASFENIEEPEERDGYGRWNHE 146
 448 ETIES-----SRAEEAKAEAP-----EEEEEAEEEGEGGEAEAE-----GEE 487

QY 147 GDD 149
 DB 488 GEE 490
 RESULT 14
 CHLD_ARATH STANDARD; PRT; 759 AA.
 ID CHLD_ARATH Q9SWE1; Q9SWE5;
 AC 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Magnesium-chelatase subunit chld, chloroplast precursor (Mg-protoporphyrin IX chelatase) (Mg-chelatase subunit D).
 GN CHLD OR ATIG08520 OR T27G7.20.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustoids II; Brassicales; Brassicaceae; Arabidopsids.
 CC NCBI_TaxID=3702;
 RN [1]
 SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016719; PubMed=11130712;
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S., White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y., Buehler E., Chen A., Chao Q., Chen K., Cheuk R.F., Chin C.W., Chung M.K., Conn L., Conway A.B., Creasy T.H., Dewar K., Dunn P., Egtu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y., Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L., Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E., Kim C.J., Koo H.B., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B., Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P., Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marzilli A., Militischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I., Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D., Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M., Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D., Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M., Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis thaliana";
 RT Nature 408:816-820(2000).
 RL [2]
 SEQUENCE OF 32-759 FROM N.A.
 RC STRAIN=cv. C24;
 RA Green J., Jensen P.E., Gibson L.C.D., Hunter C.N.;
 RT "Characterization of the magnesium protoporphyrin chelatase chld subunit from Arabidopsis thaliana cv. C24";
 RT Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Involved in chlorophyll biosynthesis; introduces a magnesium ion into protoporphyrin IX to yield Mg-protoporphyrin IX.
 CC -!- PATHWAY: Chlorophyll biosynthesis.
 CC -!- SUBCELLULAR LOCATION: Chloroplast stroma.
 CC -!- SIMILARITY: BELONGS TO THE MG-CHELATASE SUBUNITS D/I FAMILY.
 CC -!- SIMILARITY: Contains 1 VFWA domain.
 CC -!- CAUTION: Ref.1 sequence differs from that shown due to erroneous gene model prediction.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL; AC006932; AAF22895.1; ALT_SEQ.
 DR EMBL; AF083555; AAD52031.1; -;
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR000523; Mg_chelatase_chII.
 DR InterPro; IPR002035; VFW_A.

```

DR Pfam: PR01078; Mg chelatase; 1.
DR SMART; SM00382; AAA; 1.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS02344; WFA; 1.
KW Photosynthesis; Chlorophyll biosynthesis; ATP-binding; Chloroplast;
KW Transit peptide.
FT TRANSIT 1 ?
FT CHAIN 1 ?
FT DOMAIN 557 753
FT DOMAIN 411 418
FT DOMAIN 425 438
FT CONFLICT 70 70
FT CONFLICT 150 150
FT CONFLICT 270 270
FT CONFLICT 285 286
FT CONFLICT 314 314
FT CONFLICT 379 379
FT CONFLICT 426 426
FT CONFLICT 426 426
SQ SEQUENCE 759 AA; 83154 MW; A4169319E1247BD1 CRC64;

Query Match 12.4%; Score 95; DB 1; Length 759;
Best Local Similarity 24.0%; Pred. No. 1.8;
Matches 40; Conservative 21; Mismatches 42; Indels 64; Gaps 6;

QY 16 EECLEAFKQVFGSLERTRAQV---RYLKPYQ-----EEGKVSAYVL 55
Db 307 ERCNEVFRWNEBETAKTQILAREYLDVKISREQLKVLVEAVRGVGGVGHRAELYAA 366
QY 56 RLETL-----LRRAVEKRAIPRIADQVRLEQVQVAGATNLQMLWCRLRELK 101
Db 367 RVAKCLAAEGREKVTIDLRKAVELVILPRSLDTPPEQ-----Q 408
QY 102 DQGGPP-----SFLMKVIREEESEAFENESIEE-----PRE 136
Db 409 NOPPPPPPPQNSGESENEEQEEDESENEEQDQIPEE 455

RESULT 15
ID NAB2_HUMAN STANDARD; PRT; 525 AA.
AC Q15742; Q76006; Q14797;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE NGFI-A binding protein 2 (EGR-1 binding protein 2) (Melanoma-
DE associated delayed early response protein) (MADER protein).
DE NAB2 OR MADER.
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Placenta;
RX MEDLINE=96251303; PubMed=8668170;
RA Svaren J., Severson B.R., Apel E.D., Zimonjic D.B., Popescu N.C.,
RA Milbrandt J.;
RT "NAB2, a corepressor of NGFI-A (Egr-1) and Krox20, is induced by
RT proliferative and differentiative stimuli.";
RL Mol. Cell. Biol. 16:3545-3553(1996).
[2]
RN SEQUENCE FROM N.A. (ISOFORM 1).
RA Garlinger M., Johnson J.P.;
RT "Genomic organization of the Mader/NAB2 gene.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE OF 50-525 FROM N.A. (ISOFORM 1).
RX MEDLINE=96243039; PubMed=8649813;
RA Kirsch K.H., Korradi V., Johnson J.P.;
RA "Mader: a novel nuclear protein over expressed in human melanomas.";
RL Oncogene 12:963-971(1996).
[4]
RN PARTIAL SEQUENCE FROM N.A. (ISOFORM 3).
RP

```

```

RA Johnson J.P.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
[5]
RP PARTIAL SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=97271553; PubMed=9126479;
RA Svaren J., Apel E.D., Simburger K.S., Jenkins N.A., Gilbert D.J.,
RA Copeland N.A., Milbrandt J.;
RT "The Nab2 and Stat6 genes share a common transcription termination
RT region.";
RL Genomics 41:33-39(1997).
CC -!- FUNCTION: Acts as a transcriptional repressor for zinc finger
CC transcription factors EGR1 and EGR2. Isoform 2 lacks repression
CC ability (By similarity).
CC -!- SUBUNIT: Homomultimers may associate with EGR1 bound to DNA (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: NUCLEAR. ISOFORM 2 IS NOT LOCALIZED TO THE
CC NUCLEUS (BY SIMILARITY).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=Q15742-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q15742-2; Sequence=VSP_003385, VSP_003386;
CC Name=3;
CC IsoId=Q15742-3; Sequence=VSP_003387;
CC -!- TISSUE SPECIFICITY: Widely expressed at low levels. Highly
CC expressed in melanoma cell lines.
CC -!- INDUCTION: By serum and PMA stimulation.
CC -!- DOMAIN: The NAB conserved domain 1 (NCD1) interacts with EGR1
CC inhibitory domain and mediates multimerization.
CC -!- DOMAIN: The NAB conserved domain 2 (NCD2) is necessary for
CC transcriptional repression.
CC -!- SIMILARITY: BELONGS TO THE NAB FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U48361; AAC50589.1; -.
CC EMBL; AF268380; AAF72545.1; -.
CC EMBL; X70991; CAA50318.1; -.
CC EMBL; AJ011081; CAA09472.1; -.
CC Genew; HGNC:7627; NAB2.
CC MIM; 602381; -.
CC GO; GO:0003714; F:transcription co-repressor activity; TAS.
CC GO; GO:0008283; P:cell proliferation; TAS.
CC GO; GO:0007399; P:neurogenesis; TAS.
CC InterPro; IPR006989; Nab_cent.
CC InterPro; IPR006988; Nab_N.
CC Pfam; PF04904; NCD1; 1.
CC Pfam; PF04905; NCD2; 1.
CC Transcription regulation; Repressor; Alternative splicing.
FT DOMAIN 35 113 NCD1.
FT DOMAIN 267 356 NCD2.
FT DOMAIN 353 384 NECESSARY FOR NUCLEAR LOCALIZATION
FT VARSPLIC 320 322 (BY SIMILARITY).
FT FT LTI -> ASP (in isoform 2).
FT VARSPLIC 323 525 /FTId=VSP_003385.
FT FT Missing (in isoform 2).
FT VARSPLIC 426 489 /FTId=VSP_003386.
FT FT Missing (in isoform 3).
FT FT /FTId=VSP_003387.
FT FT PR -> Q (IN REF. 3 AND 4).
SQ SEQUENCE 525 AA; 56594 MW; 38CF6CFE73756F9 CRC64;

Query Match 12.3%; Score 94.5; DB 1; Length 525;
Best Local Similarity 23.5%; Pred. No. 1.3;
Matches 36; Conservative 29; Mismatches 53; Indels 35; Gaps 5;

```

Qy	4	HIVQADNPISVEECLEAFKQVFGSLSSRRTAQVRYLKPYOEGEKVS-----A	52
Db	280	HIFEMDDNDQKEEIRKYSIIYGRFDSKR-----REGQLSLHELTTINEAAQ	328
Qy	53	YVLRLETLRRRAVEKRAIPRIADQVRLEQVMAGATLN--OMLWCRRLRELKD-----	102
Db	329	FCMRDNTILLRRVELFSLSRQVARESTYLSLKGSRLLHPELGGPPLKCLKQEVGEQSHP	388
Qy	103	--QGPPPSFLELMKVIREE-EEBBAEFENESIE	132
Db	389	EIQOPPPGPESYVPPYRPSLEDSASLSGESLD	421

Search completed: September 27, 2004, 17:08:45
Job time : 6.86928 secs

B/ANK

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 27, 2004, 17:05:46 ; Search time 6.57353 Seconds
(without alignments)
2180.341 Million cell updates/sec

Title: US-10-037-860-9

Perfect score: 766

Sequence: 1 DLMHVQADNPSISVEECLE.....SIEEPEERDGYGRWNHGGDD 149

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 28366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 28366

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	102.5	13.4	554	2 JW0094	neurofilament prot
2	96	12.5	475	2 B44841	low molecular weig
3	94.5	12.3	452	2 S31927	drop9 protein - hu
4	94	12.3	548	1 QPPGL	neurofilament trip
5	93.5	12.2	544	2 S07144	neurofilament trip
6	93.5	12.2	1364	2 T14900	condensin XCAP-D2
7	93	12.1	399	2 G72253	RNA polymerase sig
8	91	11.9	657	2 S32739	filensin - chicken
9	90	11.7	543	1 QPMSL	neurofilament trip
10	90	11.7	592	2 JC7709	testis-specific ki
11	89.5	11.7	800	2 T00034	SART-1 protein - h
12	86.5	11.3	778	2 B86218	protein T27G7.20 [
13	85	11.1	387	1 S52274	transcription term
14	85	11.1	800	2 T19627	hypothetical prote
15	84.5	11.0	234	2 F70416	hypothetical prote
16	84.5	11.0	1184	2 T41515	coiled coil protei
17	84.5	11.0	1300	2 I53799	CGI protein - huma
18	84	11.0	775	2 T26186	hypothetical prote
19	83.5	10.9	775	2 H71688	DNA topoisomerase
20	83	10.8	1175	2 C35815	myosin heavy chain
21	83	10.8	1175	2 D35815	myosin heavy chain
22	83	10.8	1201	2 A35815	myosin heavy chain
23	83	10.8	1201	2 B35815	myosin heavy chain
24	83	10.8	2385	2 A32491	myosin heavy chain
25	83	10.8	2411	2 B32491	myosin heavy chain
26	82	10.7	576	2 AE3491	SSU ribosomal prot
27	82	10.7	1356	2 S32763	kinectin 1 - human
28	81.5	10.6	105	2 H72708	probable ribosomal
29	81.5	10.6	425	2 S76760	transcription init

ALIGNMENTS

RESULT 1

JW0094

neurofilament protein-L - bovine

N:Alternate names: NF-L

C:Species: Bos primigenius taurus (cattle)

C:Date: 03-Jul-1998 #sequence_revision 10-Jul-1998 #text_change 07-May-1999

C:Accession: JW0094

R:Hashimoto, R.; Nakamura, Y.; Goto, H.; Wada, Y.; Sakoda, S.; Kaibuchi, K.; Inagaki, M.; Biochem. Biophys. Res. Commun. 245, 407-411, 1998

A:Title: Domain- and site-specific phosphorylation of bovine NF-L by Rho-associated kinase A:Reference number: JW0094; MUID:98238650; PMID:9571164

A:Accession: JW0094

A:Molecule type: protein

A:Residues: 1-554 <HAS>

C:Comment: Domain- and site-specific phosphorylation by Rho-kinase regulates the assembly C:Superfamily: cytoskeletal keratin

Query Match 13.4%; Score 102.5; DB 2; Length 554;

Best Local Similarity 23.0%; Pred. No. 0.46;

Matches 41; Conservative 24; Mismatches 66; Indels 47; Gaps 6;

QY 2 LMHVQADNPSISVEECLEAFKQVFGSLRSRTAQVRLKPYQE----- 45

Db 333 LQLEEDKQNDISAMQ--DTINKLEBLRTTKSEMARYLKEYQDLLNFKMALDIEIAAYR 390

QY 46 ---EGEK-----VSAYVLRLETLERRA-----VEKRAIPRIADQVPLE 81

Db 391 KLEGEETRLSFTSVGSLTTGTTOSSQVFGKSAYGGLQTSSYLMARSFSPSYTSHVQEE 450

QY 82 QVMAGATINQMLWCRRLKDLQGGPPSPFLMLKVIREEEEAEAFENESTIEEPEERDG 139

Db 451 QIEVEETIEA---AKAEAKDE--PPSEGEAEKEKEKEAEAEAEAEAEAEAE 503

RESULT 2

B44841

low molecular weight neurofilament protein XNF-L - African clawed frog

C:Species: Xenopus laevis (African clawed frog)

C:Date: 31-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 12-Apr-1995

C:Accession: B44841

R:Charnas, L.R.; Szaro, B.G.; Gainer, H.

J. Neurosci. 12, 3010-3024, 1992

A:Title: Identification and developmental expression of a novel low molecular weight neuro A:Reference number: A44841; MUID:92356194; PMID:1494944

A:Accession: B44841

A:Molecule type: mRNA

A:Residues: 1-544 <CHA>

A:Experimental source: brain

A:Note: sequence inconsistent with the nucleotide translation

A:Note: sequence extracted from NCBI backbone (NCBIN:110225, NCBI:110226)

C:Superfamily: cytoskeletal keratin

Query Match 12.5%; Score 96; DB 2; Length 544;
Best Local Similarity 23.0%; Pred. No. 1.6;
Matches 42; Conservative 19; Mismatches 56; Indels 66; Gaps 6;
QY 9 DNPSISVVECLAFKQVFGSLRRRTAQVRLKPYOE-----EGEK 49
DB 332 DKQGEIAGQDAINKLEELRNTSEMARYLKEYODLLNVKVALDIEIAAYRKLEGE 391
QY 50 V-----SAYVLRLETLRRRAVEKRAIPRRRIADQVRLQVWAG 86
DB 332 TRLSPGVCAITSGYTQAPVGRSAYSLOSSYM-----TSRAFFYYSHVQEEQLDIE 447
QY 87 ATLNQWLRCRLRELKQDQPPFLELMKVIREEEAEAFENSTHEPERDYGWRNHE 146
DB 448 ETIES-----SRABEAKAEAP-----EEEEAEAEEGGGEAEAE-----GEE 487
QY 147 GDD 149
DB 488 GEE 490

RESULT 3
S31927
drop9 protein - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 10-Sep-1997
C:Accession: S31927
R:Kirsch, K.
submitted to the EMBL Data Library, February 1993
A:Reference number: S31927
A:Accession: S31927
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-475 <KIR>
A:Cross-references: EMBL:X70991; NID:G38459; PID:G38460

Query Match 12.3%; Score 94.5; DB 2; Length 475;
Best Local Similarity 23.5%; Pred. No. 1.8;
Matches 36; Conservative 29; Mismatches 53; Indels 35; Gaps 5;
QY 4 HIVQADNPSISVECLAFKQVFGSLRRRTAQVRLKPYOEKEKVS-----A 52
DB 230 HIFEMDNDQKEETKYSIIYGRFDSK-----REGKQLSLHELINAEAAQ 278
QY 53 YVLRLETLRRRAVEKRAIPRRRIADQVRLQVWAGATLN--OMLWCRRLRLKQ----- 102
DB 279 FCRDNTLLRLRVFLSLRSQVARESTYLSLKGRLHPEELGGPPLKLLKQEVGQSHP 338
QY 103 --QPPPSFLELMKVIREE-EEBESFENESIE 132
DB 339 EIQQPPPGPSYVPPYRPSLEDSASLSGSLD 371

RESULT 4
QFPCGL
neurofilament triplet L protein - pig
N:Alternate names: 68k neurofilament protein
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 15-Nov-1984 #sequence_revision 28-May-1986 #text_change 10-Oct-1997
C:Accession: A91337; A90973; A34569; A02963
R:Geisler, N.; Plessmann, U.; Weber, K.
FEBS Lett. 182, 475-478, 1985
A:Title: The complete amino acid sequence of the major mammalian neurofilament protein
A:Reference number: A91337; MUID:85154583; PMID:3920075
A:Accession: A91337
A:Molecule type: protein
A:Residues: 1-547 <GL>
R:Geisler, N.; Kaufmann, E.; Fischer, S.; Plessmann, U.; Weber, K.
EMBO J. 2, 1295-1302, 1983
A:Title: Neurofilament architecture combines structural principles of intermediate filaments
A:Reference number: A90973
A:Accession: A90973

A:Molecule type: protein
A:Residues: 1-82; 278-548 <GE2>
R>Note: residue 322 is either lysine or arginine
R:Gonda, Y.; Nishizawa, K.; Ando, S.; Kitamura, S.; Minoura, Y.; Nishi, Y.; Inagaki, M.
Biochem. Biophys. Res. Commun. 167, 1316-1325, 1990
A:Title: Involvement of protein kinase C in the regulation of assembly-disassembly of ne
A:Reference number: A34569; MUID:90211318; PMID:2108674
A:Accession: A34569
A:Status: preliminary
A:Molecule type: protein
A:Residues: 9-14; 23-29; 30-53 <GON>
C:Comment: Mammalian neurofilament proteins: a conserved alpha-helical region, whose hel
ke all other intermediate filament proteins: a conserved alpha-helical region, whose hel
C:Comment: The amino-terminal headpiece is basic with a high content of hydroxyamino aci
al beta turns; domain b is acidic and rich in glutamic acid and lysine residues.
C:Comment: The extra mass and high charge density that distinguish the neurofilament pro
charged scaffolding structure suitable for interaction with other neuronal components o
C:Comment: The boundaries of the domains between residues 70-92 and 399-402 are not yet
C:Superfamily: cytoskeletal keratin
C:Keywords: coiled coil; intermediate filament
F:1-70/Domain: head <HED>
F:92-123/Domain: coil 1a, alpha-helical rod <R1A>
F:137-232/Domain: coil 1b, alpha-helical rod <R1B>
F:255-399/Domain: coil 2, alpha-helical rod <RD2>
F:402-548/Domain: tail <TAI>
F:402-442/Region: tail subdomain a
F:443-548/Region: tail subdomain b

Query Match 12.3%; Score 94; DB 1; Length 548;
Best Local Similarity 22.0%; Pred. No. 2.4;
Matches 44; Conservative 24; Mismatches 70; Indels 62; Gaps 7;
QY 2 LMHIVQADNPSISVECLAFKQVFGSLRRRTAQVRLKPYOE----- 45
DB 332 LOELDKQADISAMQ--DTINKLENELRTTSEMARYLKEYODLLNVKVALDIEIAAYR 389
QY 46 ---EGEK-----VSAYVLRLETLRRRA-----VEKRAIPRRRIADQVRL 81
DB 390 KLLGEETRLSFTSGSLITGYSSQVGRSAVGLQTSSYLMSTRSPSYTSHVQEE 449
QY 82 QVMAGATLNQMLWCRRLRELKQDQPPPSFLELMKVIREE-----EEBESASF 126
DB 450 QIEVETETIA--AKAEAKDE--PPSEGEAEKEKEAEAEAEAEAEAEAEAE 504
QY 127 ENESIEPEERDYGWRNHE 146
DB 505 KEESSEAEKEEGEGEGEE 524

RESULT 5
S07144
neurofilament triplet L protein - human
N:Alternate names: neurofilament light polypeptide (68k)
N:Contains: Glu-50 brain peptide
C:Species: Homo sapiens (man)
C>Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 21-Jul-2000
R:Julien, J.P.; Grosfeld, F.; Yazdankhah, K.; Flavell, D.; Meijer, D.; Mushynski, W.
Biochim. Biophys. Acta 909, 10-20, 1997
A:Title: The structure of a human neurofilament gene (NF-L): a unique exon-intron organ
A:Reference number: S07144; MUID:87214213; PMID:3034332
A:Accession: S07144
A:Molecule type: DNA
A:Residues: 1-544 <JUL>
A:Cross-references: EMBL:X05608; NID:G1495072; PID:CAA29097.1; PID:G1279504
R>Note: the authors translated the codon ATG for residue 366 as Asn
R:Postelov, V.A.; Pospelova, T.V.; Julien, J.P.
Cell Growth Differ. 5, 187-196, 1994
A:Title: AP-1 and Krox-24 transcription factors activate the neurofilament light gene p
A:Reference number: 152832; MUID:94235564; PMID:8180132
A:Accession: 152832
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
A:Residues: 1-10 <POS>
A:Cross-references: GB:S70309; NID:g547176; PIDN:AAD14057.1; PID:g4261757
R;Nomata, Y.; Watanabe, T.; Wada, H.
J. Biochem. 93, 825-831, 1993
A:Title: Highly acidic proteins from human brain: purification and properties of Glu-50
A:Reference number: A60703; MUID:83265667; PMID:6135695
A:Accession: A60703
A:Molecule type: protein
A:Residues: 469-472, 'D', 474 <NOM>
A:Experimental source: Glu-50 brain peptide
A:Note: This acidic protein is named for its greater than fifty per cent glutamic acid c
C:Genetics:
A:Gene: GDB:NEFL; NFL
A:Cross-references: GDB:120227; OMIM:162280
A:Map position: 9p21-8p21
A:Introns: 349/3; 391/2; 498/1
C:Superfamily: cytoskeletal keratin
C:Keywords: brain; coiled coil; intermediate filament
F:469-544/Product: Glu-50 peptide #status predicted <B50>

Query Match 12.2%; Score 93.5; DB 2; Length 544;
Best Local Similarity 23.1%; Pred. No. 2.6;
Matches 45; Conservative 24; Mismatches 71; Indels 55; Gaps 8;

QY 2 LMHIQADNPSISVEECLEAFKQVFGSLESRTAQVRYLKPQF-----45
Db 334 LQELDKQNDISAMQ--DTINKLENELRTTKSEMARYLKEYQDLLNVKQWALDIEIAAYR 391
QY 46 ---EGEX-----VSAYVLRLETLRRA-----VEKRAIRIADQVRLE 81
Db 392 KLLGEETRLSFTSGVITSGYSSQVFGESAGGLQTSYLMSTSFPSYITSHVQEE 451
QY 82 QVWAGATLQMLWCRRLRELKQDQPPPSFLELMKVIREE-EEEEASFENESIEE-----P 134
Db 452 QTEVEETIEA---SKAEAKDEPSEGEAEKEEKEEAEAEAEAEAEAEAEAEAEAE 508
QY 135 EERDGYGRWNHGDD 149
Db 509 EEEGEGE---EEGEE 520

RESULT 6
T14900
condensin XCAP-D2 chain - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 25-Aug-2003
C:Accession: T14900
R;Kimura, K.; Hirano, M.; Kobayashi, R.; Hirano, T.
Science 282, 487-490, 1998
A:Title: Phosphorylation and activation of 13S condensin by cdc2 in vitro.
A:Reference number: Z18257; MUID:98447791; PMID:9774278
A:Accession: T14900
A:Status: preliminary;
A:Molecule type: mRNA
A:Residues: 1-1364 <KIM>
A:Cross-references: EMBL:AF067969; NID:g3764086; PID:g3764087; PIDN:AAC64359.1
C:Superfamily: chromosome condensation complex Condensin, subunit D2

Query Match 12.2%; Score 93.5; DB 2; Length 1364;
Best Local Similarity 22.3%; Pred. No. 7.3;
Matches 35; Conservative 26; Mismatches 45; Indels 51; Gaps 6;

QY 30 ESRTAQVRYLKPQEGEKVSAVLR-LBTLRRAVEKRAIP-RRIADQVRLE-----81
Db 352 EAEKSSRDQFLDTLQELHVDWTVRSCVQIYNRIQVEXALPLSRFSQSVTLVWGRFLD 411
QY 82 -----QVWAGATLNQMLACRL-----RELKD-----QGPP-----106
Db 412 KSNVCKNAIQLLASFANNPFTCKLSSVDLKVPLEKETKLLKEMREKYQGPKPVVWISP 471
QY 107 -----PSFLELMKVIREEEAEASFENESIEEPE 135

Db 472 EEEWEAMLPVLEAFKILQQESKEEDIEETEEISSQ 508

RESULT 7
G72253
RNA polymerase sigma-A factor - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: G72253
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinm, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A:Reference number: A72200; MUID:99287316; PMID:10360571
A:Accession: G72253
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-399 <ARN>
A:Cross-references: GB:AE001796; GB:AE000512; NID:g4982004; PIDN:AAD36519.1; PID:g498201
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM1451
C:Superfamily: transcription initiation factor sigma 43; transcription initiation factor
F:161-397/Domain: transcription initiation factor sigma katF homology <KTF>

Query Match 12.1%; Score 93; DB 2; Length 399;
Best Local Similarity 22.4%; Pred. No. 2;
Matches 44; Conservative 31; Mismatches 63; Indels 58; Gaps 6;

QY 1 DLMIHQADNPSI---SVEECLEAFKQVFGSLESRTAQVRYLKPQEGEKVSAVLR 56
Db 101 ELBELLEKESPEIHDSNVSDSIKMYLKEIGKIPLLTPAQERELARRAQMGDKKAKELI 160
QY 57 LETL-----LRRAVEK-----RAIP 71
Db 161 TSNLRVSTAKRYMGRLSFDQLIQEGNIGLLKAVEKFDWRKGYKSTVATWIRQAIT 220
QY 72 RRIADQVRLEQVWAG--ATLNQMLWCRRLRELKQDQPPPSFLELMKVIREEEEE--EA 124
Db 221 RA1ADQARTIRIPVHMVETINKLNRRLRYVYQKHGEPSTIEELAKMGKPEKIKELEA 280
QY 125 SFENESIEEP--BERD 138
Db 281 AKETISLESPIGEDED 296

RESULT 8
S32739
filensin - chicken
C:Species: Gallus gallus (chicken)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 10-Sep-1997
C:Accession: S32739
R;Remington, S.G.
submitted to the EMBL Data Library, March 1993
A:Description: Chicken filensin: a lens fiber cell protein exhibits sequence similarity t
A:Reference number: S32739
A:Accession: S32739
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-657 <REM>
A:Cross-references: EMBL:X72873; NID:g296208; PID:g296209

Query Match 11.9%; Score 91; DB 2; Length 657;
Best Local Similarity 25.7%; Pred. No. 5.2;
Matches 37; Conservative 29; Mismatches 50; Indels 28; Gaps 6;

QY 20 EAFKQVFGSLESRTAQVRYLKPQEGEKVSAVLR-----ETLLRRAVEKRAIPRR 73
Db 19 DAYDELGRSPFSDSLAQAOGLNELNRFASYNRVARVLEQRNLTILRKQLETF---OR 75
QY 74 IADQVRLEQVWAG--ATLNQMLWCRRLRELKQDQPPPSFLELMKVIREE-----BEEAS 125

Db 76 MDELVLDEAFAGQIEFNQ-----RMRELAS-----DRAKLEREKDAQRMLEDYHNK 124

QY 126 FNESEIEPEERDGYGRWNHGGD 149
: || : : : || :
Db 125 YRNEYEQQKLTLERLNKEADE 148

RESULT 9

QFMSL

neurofilament triplet L protein - mouse

N:Alternate names: 68K neurofilament protein; NF-L(low) protein; type IV IF protein

C:Species: Mus musculus (house mouse)

C:Date: 31-Mar-1988 #sequence revision 30-Sep-1993 #text_change 22-Jun-1999

C:Accession: A25227; A26562; A3772; A41012; 155316

R:Lewis, S.A.; Cowan, N.J.

Mol. Cell. Biol. 6, 1529-1534, 1986

A:Title: Anomalous placement of introns in a member of the intermediate filament multigene family

A:Reference number: A25227; MUID:87064433; PMID:3785173

A:Accession: A25227

A:Molecule type: DNA

A:Residues: 1-543 <LEW>

A:Cross-references: GB:M13016; NID:G200023; PIDN:AAA39810.1; PID:G387492

A:Note: The authors translated the codon GGC for residue 5 as Ala, ACA for residue 88 as 1 as Glu

R:Lewis, S.A.; Cowan, N.J.

J. Cell Biol. 100, 843-850, 1985

A:Title: Genetics, evolution, and expression of the 68,000-mol-wt neurofilament protein: A reference number: A26562; MUID:85131334; PMID:3919033

A:Accession: A26562

A:Molecule type: mRNA

A:Residues: 242-543 <LB2>

A:Cross-references: GB:X02165

A:Experimental source: brain

R:Julien, J.P.; Meyer, D.; Flavell, D.; Hurst, J.; Grosveld, F.

Brain Res. Mol. Brain Res. 1, 243-250, 1986

A:Title: Cloning and developmental expression of the murine neurofilament gene family.

A:Reference number: A43772

A:Accession: A43772

A:Molecule type: mRNA

A:Residues: 1-5, 'Y', 7-8, 'Y', 10-64, 'M', 66-72, 'L', 74-98, 'D', 100-194, 'R', 196-202, 204-239, 'Y'

A:Cross-references: GB:M20480; NID:G200037; PIDN:AAA39814.1; PID:G200038

A:Note: The authors translated the codon CGC for residue 195 as Ala

R:Sihaq, R.K.; Nixon, R.A.

J. Biol. Chem. 266, 18861-18867, 1991

A:Title: Identification of Ser-55 as a major protein kinase A phosphorylation site on the regulatory subunit of the rat brain neurofilament protein

A:Reference number: A41012; MUID:92011653; PMID:1717455

A:Accession: A41012

A:Molecule type: protein

A:Residues: 52-57 <SIH>

R:Nakahira, K.; Ikenaka, K.; Wada, K.; Tamura, T.

J. Biol. Chem. 265, 19786-19791, 1990

A:Title: Structure of the 68-kDa neurofilament gene and regulation of its expression.

A:Reference number: I55316; MUID:91060592; PMID:2246261

A:Accession: I55316

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-5, 'Y', 7-8, 'Y', 10-28 <RES>

A:Cross-references: GB:M55423; NID:G200027; PIDN:AAA39812.1; PID:G554245

A:Comment: This is the most abundant of the three neurofilament proteins and, as the other two, is a major component of the cytoskeleton

C:Genetics:

A:Introns: 349/3; 391/2; 498/1

C:Superfamily: cytoskeletal keratin

C:Keywords: coiled coil; intermediate filament

F:2-72/Domain: head <HED>

F:94-125/Domain: coil 1a, alpha-helical rod #status predicted <R1A>

F:126-138/Region: linker 1

F:139-234/Domain: coil 1b, alpha-helical rod #status predicted <R1B>

F:235-256/Region: linker 12

F:257-272/Domain: coil 2a, alpha-helical rod #status predicted <R2A>

F:273-281/Region: linker 2

F:282-401/Domain: coil 2b, alpha-helical rod #status predicted <R2B>

F:404-543/Domain: tail <TAI>

F:404-444/Region: tail subdomain a

F:445-543/Region: tail subdomain b

Query Match 11.7%; Score 90; DB 1; Length 543;

Best Local Similarity 22.4%; Pred. No. 5.1;

Matches 44; Conservative 26; Mismatches 66; Indels 60; Gaps 8;

C;Superfamily: Rhodobacter sphaeroides protoporphyrin IX magnesium chelatase bchD

```

16  ECLEAFKQVFGSLESRRTAQV---RYLKPKYQ-----EEGKVSAYVL 55
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
307  ERCNEVFARMNEETETAKTQIILAREYLKQVKVRSRQLKYLVEARGVQGVQRAELVAA 366
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

56  RLETL-----LRRAVEKRAIPRRIDQVRLQVMAGATINQMLWCRLRELK 101
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
367  RVAKCLAAIEGREKVTIDLSKAVELVLPSSSIDETPPEO-----O 408
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```

S52274
transcription termination-antitermination factor nusA - *Thermus aquaticus*
C;Species: *Thermus aquaticus*
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 15-Oct-1999
C;Accession: S52274
R;Vornlocher, H.; Sprinzl, M.

RESULT 15

RESULT 15

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 27, 2004, 17:10:17 ; Search time 43.3366 Seconds
(without alignments)
1105.584 Million cell updates/sec

Title: US-10-037-860-9
 Perfect score: 766
 Sequence: 1 DLWHIVQADNPSISVECLE.....SIEPERDGYGRNHEGDD 149

Scoring table: BLOSUM62
Gapop 10.0 ; Gapext 0.5

Searched: 1349238 seqs, 321558718 residues

Total number of hits satisfying chosen parameters: 1349238

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 s

```

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pcp.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp.*
3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pcp.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS0_PUBCOMB.pcp.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pcp.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pcp.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pcp.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pcp.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pcp.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pcp.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pcp.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pcp.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp.*

```

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	766	100.0	149	13	US-10-037-860-9		Sequence 9, Appli
2	755	98.6	283	13	US-10-037-860-11		Sequence 11, Appl
3	323	42.2	463	13	US-10-037-860-13		Sequence 13, Appl
4	295.5	38.6	353	9	US-09-965-529-7		Sequence 7, Appl
5	295.5	38.6	353	10	US-09-969-680A-7		Sequence 7, Appl
6	287	37.5	452	15	US-10-048-765A-2385		Sequence 2385, Ap
7	277	36.2	399	15	US-10-094-749-1978		Sequence 1978, Ap
8	271.5	35.4	351	9	US-09-965-529-1		Sequence 1, Appli
9	271.5	35.4	351	10	US-09-969-680A-1		Sequence 1, Appl
10	271.5	35.4	351	12	US-09-040-014A-16		Sequence 16, Appl
11	271.5	35.4	351	15	US-10-341-434-10		Sequence 10, Appl
12	256.5	33.5	329	13	US-10-037-860-4		Sequence 4, Appl
13	255	33.3	318	12	US-09-804-014A-40		Sequence 40, Appl
14	247.5	32.3	403	15	US-10-094-466-38		Sequence 38, Appl
15	240	31.3	337	12	US-10-296-115-1208		Sequence 1208, Ap

16	235	30.7	321	12	US-09-804-014A-39	Sequence 39, Appl
17	225	29.4	312	12	US-09-804-014A-73	Sequence 73, Appl
18	225	29.4	312	12	US-09-804-014A-74	Sequence 74, Appl
19	97.5	12.7	218	15	US-10-094-749-2881	Sequence 2881, Ap
20	96.5	12.6	542	12	US-10-205-331-57	Sequence 57, Appl
21	95.5	12.5	407	16	US-10-755-889-122	Sequence 122, App
22	93.5	12.2	620	16	US-10-437-963-158544	Sequence 158544,
23	91.5	11.9	346	15	US-10-310-154-A48	Sequence 448, App
24	90	11.7	592	12	US-10-416-477-8	Sequence 8, Appli
25	90	11.7	592	12	US-10-438-339-8	Sequence 8, Appli
26	89	11.6	1474	16	US-10-437-963-187531	Sequence 187531,
27	87.5	11.4	2552	16	US-10-437-963-129734	Sequence 129734,
28	87	11.4	545	9	US-09-908-398B-3	Sequence 4, Appli
29	87	11.4	545	16	US-10-775-649-4	Sequence 4, Appli
30	87	11.4	545	16	US-10-775-627-4	Sequence 4, Appli
31	84.5	11.0	1300	16	US-10-408-765A-257	Sequence 257, App
32	83	10.8	882	14	US-10-298-417-2	Sequence 2, Appli
33	83	10.8	1687	16	US-10-437-963-187527	Sequence 187527,
34	83	10.8	1708	16	US-10-437-963-187533	Sequence 187533,
35	82.5	10.8	384	16	US-10-437-963-182739	Sequence 182739,
36	82.5	10.8	435	9	US-09-866-582-33	Sequence 33, Appl
37	82.5	10.8	549	16	US-10-437-963-156505	Sequence 156505,
38	82.5	10.8	817	12	US-10-363-616-A47	Sequence 447, App
39	82.5	10.8	996	16	US-10-380-492A-4	Sequence 4, Appli
40	82.5	10.8	1002	16	US-10-380-492A-2	Sequence 2, Appli
41	82.5	10.8	1009	14	US-10-290-544-2	Sequence 2, Appli
42	82.5	10.8	1009	14	US-10-290-544-5	Sequence 5, Appli
43	82.5	10.8	1023	15	US-10-108-260A-3123	Sequence 3123, Ap
44	82.5	10.8	1023	16	US-10-380-492A-8	Sequence 8, Appli
45	82	10.7	715	16	US-10-408-765A-2096	Sequence 2096, AP

ALIGNMENTS

```

RESULT 1
US-10-037-860-9
; Sequence 9, Application US/10037860
; Publication No. US20020123114A1
; GENERAL INFORMATION:
; APPLICANT: Jerome B. Posner
; APPLICANT: Josep O. Salmau
; APPLICANT: Myrna R. Rosenfeld
; TITLE OF INVENTION: Ma FAMILY POLYPEPTIDES AND ANTI-Ma
; TITLE OF INVENTION: ANTIBODIES
; FILE REFERENCE: 2581.1004-004
; CURRENT APPLICATION NUMBER: US/10/037,860
; CURRENT FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 09/189,527
; PRIOR FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 149
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-037-860-9

```

```
RESULT 2
US-10-037-860-11
; Sequence 11, Application US/10037860
; Publication No. US20020123114A1
; GENERAL INFORMATION:
; APPLICANT: Jerome B. Posner
; APPLICANT: Josep O. Dalmau
; APPLICANT: Myrna R. Rosenfeld
; TITLE OF INVENTION: Ma FAMILY POLYPEPTIDES AND ANTI-Ma
; TITLE OF INVENTION: ANTIBODIES
; FILE REFERENCE: 2581.1004-004
; CURRENT APPLICATION NUMBER: US/10/037,860
; CURRENT FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 09/189,527
; PRIOR FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 283
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-037-860-11

Query Match      98.6%; Score 755; DB 13; Length 283;
Best Local Similarity 98.7%; Pred. No. 9.2e-70;
Matches 147; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DLMHIVQADNPISISVEECLEAFKQVFGSLESRTAQVRYLKPYQEGEKVSAYVLRLETL 60
DB 135 DLMHIVQADNPISISVEECLEAFKQVFGSLESRTAQVRYLKPYQEGEKVSAYVLRLETL 194

QY 61 LRRAVEKRAIPRIADQVRLQVMAGATLQMLWCRLRELKDQGGPPSPFLMLKVIREE 120
DB 195 LKAVYKRAIPRIADQVRLQVMAGATLQMLWCRLRELKDQGGPPSPFLMLKVIREE 254

QY 121 EREASFENESIEPEPERDGYGRWNHGGD 149
DB 255 EREASFENESIEPEPERDGYGRWNHGGD 283

RESULT 3
US-10-037-860-13
; Sequence 13, Application US/10037860
; Publication No. US20020123114A1
; GENERAL INFORMATION:
; APPLICANT: Jerome B. Posner
; APPLICANT: Josep O. Dalmau
; APPLICANT: Myrna R. Rosenfeld
; TITLE OF INVENTION: Ma FAMILY POLYPEPTIDES AND ANTI-Ma
; TITLE OF INVENTION: ANTIBODIES
; FILE REFERENCE: 2581.1004-004
; CURRENT APPLICATION NUMBER: US/10/037,860
; CURRENT FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 09/189,527
; PRIOR FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 463
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-037-860-13

Query Match      42.2%; Score 323; DB 13; Length 463;
Best Local Similarity 52.7%; Pred. No. 9.6e-25;
Matches 68; Conservative 26; Mismatches 33; Indels 2; Gaps 1;

QY 6 VOADNPISVEECLEAFKQVFGSLESRTAQVRYLKPYQEGEKVSAYVLRLETLRRAV 65
DB 220 LRASNASIVVECLAAQQVFGVESHKTAQVKLKAYQEGEKVSFVLRLEPLQRAV 279
```

```
QY 66 EKRAIPRIADQVRLQVMAGATLQMLWCRLRELKDQGGPPSPFLMLKVIREEEAS 125
DB 280 ENNVSRNVNQTRLKRVLSGATLPDKLRDKUMKQRKPPGCFALAVKLRREEWEAT 339

QY 126 F--ENESIE 132
DB 340 LGPDRESLE 348

RESULT 4
US-09-965-529-7
; Sequence 7, Application US/09965529
; Publication No. US20020182671A1
; GENERAL INFORMATION:
; APPLICANT: LAL, Preeti
; APPLICANT: YUE, Henry
; APPLICANT: TANG, Y. Tom
; APPLICANT: BANDMAN, Olga
; APPLICANT: BURFORD, Neil
; APPLICANT: AZIMZAI, Valda
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LU, Dying Aina M.
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0731 USA
; CURRENT APPLICATION NUMBER: US/09/965,529
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/149,641; 60/164,203; PCT/US00/22315
; PRIOR FILING DATE: 1999-08-17; 1999-11-09; 2000-08-14
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PERL Program
; SEQ ID NO 7
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020182671A1 2483172CD1
US-09-965-529-7

Query Match      38.6%; Score 295.5; DB 9; Length 353;
Best Local Similarity 46.3%; Pred. No. 4.8e-22;
Matches 62; Conservative 29; Mismatches 36; Indels 7; Gaps 2;

QY 1 DLMHIVQADNPISISVEECLEAFKQVFGSLESRTAQVRYLKPYQEGEKVSAYVLRLETL 60
DB 212 DVIRILKSNPAITTAECLEKALEQVFGSVESRDQIKFLNTYQNPGEKLSAYVIRLEPL 271

QY 61 LRRAVEKRAIPRIADQVRLQVMAGATLQMLWCRLRELKDQGGPPSPFLMLKVI 116
DB 272 LQKVVEKGAIDKDNVQARLEQVIAGANHSQAIRQLML---TGAGEGAPNLFQLLVQI 328

QY 117 REEEEEEASFENES 130
DB 329 REEEAKEEEEAEEA 342

RESULT 5
US-09-969-680A-7
; Sequence 7, Application US/09969680A
; Publication No. US20030124649A1
; GENERAL INFORMATION:
; APPLICANT: LAL, Preeti; YUE, Henry
; APPLICANT: TANG, Y. Tom; BANDMAN, Olga
; APPLICANT: BURFORD, Neil; AZIMZAI, Valda
; APPLICANT: BAUGHN, Mariah R.; LU, Dying Aina M.
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0731-1 USA
; CURRENT APPLICATION NUMBER: US/09/969,680A
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US00/22315
; PRIOR FILING DATE: 2000-08-14
```



```
; PRIOR APPLICATION NUMBER: 60/149,641
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/164,203
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PERL Program
; SEQ ID NO 7
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: incyte ID No. US20030124649A1 2483172CD1
US-09-969-680A-7

Query Match      38.6%; Score 295.5; DB 10; Length 353;
Best Local Similarity 46.3%; Pred. No. 4.8e-22;
Matches 62; Conservative 29; Mismatches 36; Indels 7; Gaps 2;

Qy 1 LMHIVQADNPISVVEECLEAFKQVFGSLESRRTAQVRYLKPQYEEGKVSAYVRLLETL 60
Db 212 DVIRILKSNPAITAECLKALEQVFGSVSSRDAQIKFLNTYQMPGEKLSAYVIRLEPL 271

Qy 61 LRAVEKRAIPRRITADQVRLQVWAGA---TLNQMLWCRLRELKDGPPPSFLELMKVI 116
Db 272 LQKVEKAGTDKUNVQARLEQVIAGNHSGAIRQLWL---TGAGGPAENLQQLIVQI 328

Qy 117 REEEERASFENES 130
Db 329 REEAKKEEEAE 342

RESULT 6
US-10-408-765A-2385
; Sequence 2385, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2385
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2385

Query Match      37.5%; Score 287; DB 16; Length 452;
Best Local Similarity 43.8%; Pred. No. 5e-21;
Matches 63; Conservative 30; Mismatches 43; Indels 8; Gaps 2;

Qy 2 LMHIVQADNPISVVEECLEAFKQVFGSLESRRTAQVRYLKPQYEEGKVSAYVRLLETL 61
Db 215 IMRVLAQNDISITVEQCLDALKIFGDKEDFRASQFRPLQTSPIKIGEKVSTFLRLLEPL 274

Qy 62 RRAVEKRAIPRRITADQVRLQVWAGATLQMLWCRLRELKDGPPPSFLELMKVIREEE 121
Db 275 QKAVEKPLSVSRDMDRLKHLARVAMTPAIRKLELLDQRCPPNPFLELMKLIREEE 334

Qy 122 ---EEASFENESIEEPEERDGYR 142
Db 335 WENTEAVMKNK-----EKPSGRGR 353
```

```
RESULT 7
US-10-094-749-1978
; Sequence 1978, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOTYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1978
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-1978

Query Match      36.2%; Score 277; DB 15; Length 399;
Best Local Similarity 43.5%; Pred. No. 4.6e-20;
Matches 60; Conservative 30; Mismatches 48; Indels 0; Gaps 0;

Qy 2 LMHIVQADNPISVVEECLEAFKQVFGSLESRRTAQVRYLKPQYEEGKVSAYVRLLETL 61
Db 208 LVHALLAENPARTAQDCILAALAQVFGDNEQAATIRVKCLTAQQQSGERLSAFVLRLEVL 267

Qy 62 RRAVEKRAIPRRITADQVRLQVWAGATLQMLWCRLRELKDGPPPSFLELMKVIREEE 121
Db 268 QKAVEKALARASADRVRLQMLTRAHLTEPLDEALRLRMAGRSPSFLEMLGLVRESEA 327

Qy 122 EEASFENESIEEPEERDQ 139
Db 328 WEASLARSVRAQTQEGAG 345

RESULT 8
US-09-965-529-1
; Sequence 1, Application US/09965529
; Publication No. US20020182671A1
; GENERAL INFORMATION:
; APPLICANT: LAL, Preeti
; APPLICANT: YUE, Henry
; APPLICANT: TANG, Y. Tom
; APPLICANT: BANDMAN, Olga
; APPLICANT: BURFORD, Neil
; APPLICANT: AZIMZAI, Yaida
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0731 USA
```

```
; CURRENT APPLICATION NUMBER: US/09/965,529
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/149,641; 60/164,203; PCT/US00/22315
; PRIOR FILING DATE: 1999-08-17; 1999-11-09; 2000-08-14
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020182671A1 112301CD1
US-09-965-529-1

Query Match      35.4%; Score 271.5; DB 9; Length 351;
Best Local Similarity 44.6%; Pred. No. 1.5e-19;
Matches 58; Conservative 32; Mismatches 29; Indels 11; Gaps 3;

QY      1  DLMHIVQADNPISIVEECLAEKQVFGSLERRTAQVRLKPYQEGEKVSAYVLRLETL 60
Db      1  DLMHIVQADNPISIVEECLAEKQVFGSLERRTAQVRLKPYQEGEKVSAYVLRLETL 60
; :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
216  DVIRVLKINNPLITVDECLQALEEVFGVTDNPRELQVKYLTYYQDDEKLSAYVLRLEPL 275
; :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY      61  LRRAVEKRAIPRIADQVRLQVMAGA---TLNQMLWCRRLKDKGPPPSFLELMKVIR 117
Db      61  LRRAVEKRAIPRIADQVRLQVMAGA---TLNQMLWCRRLKDKGPPPSFLELMKVIR 117
; :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
276  LQKLVQGAIERDAVNQARLDQVIAGAVHKTIIRREL-----NLPEDGAPGFLQLLVLIK 330
Db      276  LQKLVQGAIERDAVNQARLDQVIAGAVHKTIIRREL-----NLPEDGAPGFLQLLVLIK 330
; :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY      118  E---EEEEE 124
Db      118  E---EEEEE 124
; :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
331  DYEAEEEEEA 340
Db      331  DYEAEEEEEA 340
; :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::

RESULT 9
US-09-969-680A-1
; Sequence 1, Application US/09969680A
; Publication No. US20030124649A1
; GENERAL INFORMATION:
; APPLICANT: TANG, Y. Tom; BANDMAN, Olga
; APPLICANT: BURFORD, Neil; AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.; LU, Dyung Aina M.
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0731-1 USA
; CURRENT APPLICATION NUMBER: US/09/969,680A
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US00/22315
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/149,641
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/164,203
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030124649A1 112301CD1
US-09-969-680A-1

Query Match      35.4%; Score 271.5; DB 10; Length 351;
Best Local Similarity 44.6%; Pred. No. 1.5e-19;
Matches 58; Conservative 32; Mismatches 29; Indels 11; Gaps 3;

QY      1  DLMHIVQADNPISIVEECLAEKQVFGSLERRTAQVRLKPYQEGEKVSAYVLRLETL 60
Db      1  DLMHIVQADNPISIVEECLAEKQVFGSLERRTAQVRLKPYQEGEKVSAYVLRLETL 60
; :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
216  DVIRVLKINNPLITVDECLQALEEVFGVTDNPRELQVKYLTYYQDDEKLSAYVLRLEPL 275
; :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY      61  LRRAVEKRAIPRIADQVRLQVMAGA---TLNQMLWCRRLKDKGPPPSFLELMKVIR 117
Db      61  LRRAVEKRAIPRIADQVRLQVMAGA---TLNQMLWCRRLKDKGPPPSFLELMKVIR 117
; :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
276  LQKLVQGAIERDAVNQARLDQVIAGAVHKTIIRREL-----NLPEDGAPGFLQLLVLIK 330
Db      276  LQKLVQGAIERDAVNQARLDQVIAGAVHKTIIRREL-----NLPEDGAPGFLQLLVLIK 330
; :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY      118  E---EEEEE 124
Db      118  E---EEEEE 124
; :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
331  DYEAEEEEEA 340
Db      331  DYEAEEEEEA 340
; :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::

RESULT 10
US-09-804-014A-16
; Sequence 16, Application US/09804014A
; Publication No. US20030064489A1
; GENERAL INFORMATION:
; APPLICANT: Li, Li
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Vernet, Corine
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard
; APPLICANT: Spaderina, Steven
; APPLICANT: Majumder, Kumud
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-721 US
; CURRENT APPLICATION NUMBER: US/09/804,014A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/188,316
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/188,277
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/189,139
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/189,140
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/190,401
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/190,231
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
; GENERAL INFORMATION:
US-09-804-014A-16

Query Match      35.4%; Score 271.5; DB 12; Length 351;
Best Local Similarity 44.6%; Pred. No. 1.5e-19;
Matches 58; Conservative 32; Mismatches 29; Indels 11; Gaps 3;

QY      1  DLMHIVQADNPISIVEECLAEKQVFGSLERRTAQVRLKPYQEGEKVSAYVLRLETL 60
Db      1  DLMHIVQADNPISIVEECLAEKQVFGSLERRTAQVRLKPYQEGEKVSAYVLRLETL 60
; :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
216  DVIRVLKINNPLITVDECLQALEEVFGVTDNPRELQVKYLTYYQDDEKLSAYVLRLEPL 275
; :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY      61  LRRAVEKRAIPRIADQVRLQVMAGA---TLNQMLWCRRLKDKGPPPSFLELMKVIR 117
Db      61  LRRAVEKRAIPRIADQVRLQVMAGA---TLNQMLWCRRLKDKGPPPSFLELMKVIR 117
; :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
276  LQKLVQGAIERDAVNQARLDQVIAGAVHKTIIRREL-----NLPEDGAPGFLQLLVLIK 330
Db      276  LQKLVQGAIERDAVNQARLDQVIAGAVHKTIIRREL-----NLPEDGAPGFLQLLVLIK 330
; :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY      118  E---EEEEE 124
Db      118  E---EEEEE 124
; :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
331  DYEAEEEEEA 340
Db      331  DYEAEEEEEA 340
; :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::

RESULT 11
US-10-341-434-10
; Sequence 10, Application US/10341434
; Publication No. US20030215835A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Differentially Regulated Prostate Cancer Genes
; FILE REFERENCE: 9U 204 205 R1
; CURRENT APPLICATION NUMBER: US/10/341,434
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/348,164
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 60/348,119
```

```

; PRIOR FILING DATE: 2002-01-15
;
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: PatentIn version 3.1.1
; SEQ ID NO 10
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-341-434-10

```

Query Match		35.4%;	Score	271.5;	DB	15;	Length	351;
Best Local Similarity		44.6%;	Pred.	No.	1.5e-19;			
Matches		58;	Conservative	32;	Mismatches	29;	Indels	11; Gaps 3;
QY	1	DLMHIIQAADNPSIVSECLFAFKOVFGSLERRTAQVRYLKPYOEGEKUASAYVLRLTFL	60	:	:	:	:	:
Db	216	DVRILVKINPLITVDCLQALBEEVGVTGNPRELQVKYLTTYQKDEKUASAYVLRLLEP	275	:	:	:	:	:
QY	61	LRRAVEKRATPRRIARDQVRLEQNMGAGA---TLNQMLWCRLLRELKDQGPPSFLEIMKVIR	117	:	:	:	:	:
Db	276	LQLKVQGALEDRAVNQAARLDQTAVGAVHTIRREL-----NLPEDPAGFQLIQLLVLIK	330	:	:	:	:	:
QY	118	E----EEEEEA	124	:	:	:	:	:
Db	331	DYEAAEIEEEA	340	:	:	:	:	:

```

RESULT 12
US-10-037-860-4
; Sequence 4, Application US/10037860
; Publication No. US20020123114A1
; GENERAL INFORMATION:
; APPLICANT: Jerome B. Posner
; APPLICANT: Joseph O. Dalmiau
; APPLICANT: Myrna R. Rosenfeld
; TITLE OF INVENTION: Ma FAMILY POLYPEPTIDES AND ANTI-Ma
; TITLE OF INVENTION: ANTIBODIES
; FILE REFERENCE: 2581.1004-004
; CURRENT APPLICATION NUMBER: US/10/037,860
; CURRENT FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 09/189,527
; PRIOR FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 329
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-037-860-4

```

Query Match	33.5%	Score 256.5;	DB 13;	Length 329;
Best Local Similarity	46.6%;	Pred. No. 4.8e-10;		
Matches	54;	Conservative 25;	Mismatches 30;	Indels 7; Gaps 2
QY	1	DLMHIVQADNPSISVEBCELEAFKQVFGSLSSRRTAQVYLKPQEERKEKVSAVVRLLETL	60	
	:	:: :::: :::	: : : : : :	
Db	212	DVIRILKSNPAITTAELCKALEQWFGSVSSRDAQIKFLNTYQNPGEXLSAYVIRLEPL	271	
	:	:: :::: :::	: : : : : :	
QY	61	LRAVEKEAIPRRTADQVRLEQNWAGA-----TLNQMLWCBLRELKCOGPPPPFLEL	112	
	:	:: :::: :::	: : : : : :	
db	272	LQKVVEKAIDKDNNVQAARLEVQTAGAHSGAIRRLQLW---TCAGEGGPKPPLSV	324	
	:	:: :::: :::	: : : : : :	

RESULT 13
US-09-804-014A-40
; Sequence 40, Application US09804014A
; Publication No. US2003006449A1
; GENERAL INFORMATION:
; APPLICANT: Li, Li
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Vernet, Corine
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard
; APPLICANT: Spaderna, Steven

```

; APPLICANT: Majumder, Kumud
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-721 US
; CURRENT APPLICATION NUMBER: US/09/804,014A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/188,316
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/188,277
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/189,139
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/189,140
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/190,401
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/190,231
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (20)
; OTHER INFORMATION: Wherein Xaa is any amino acid as defined in the
; OTHER INFORMATION: specification
US-09-804-014A-40

```

```

Query Match      33.3%; Score 255; DB 12; Length 318;
Best Local Similarity 51.0%; Pred. No. 6.5e-18;
Matches 50; Conservative 23; Mismatches 21; Indels 4; Gaps 1

Qy      1  DLMHIIVQADNPSISVEECLEAFQVFGSLSRRTAQVRYLKPYQEEGKVSAYVIRLET 60
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      212 DVIRILKSNPAITABCLKALEQVFGSVSSRDAQIKFLNTYQNPGEKLSAYVIRLEP 271

Qy      61  LRRAVEKGALPRIADQVRLEQVWAGA-----TLNQMLW 94
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      272 LQKVVEKGADKQNVNQAARLEOVIAAGHSGATRRQLW 309

```

RESULT 14
US-10-094-466-38
; Sequence 38, Application US/10094466
; Publication No. US20030203363A1
; GENERAL INFORMATION:
; APPLICANT: Spytex et al.
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM
; TITLE OF INVENTION: AND METHODS OF USING
; TITLE OF INVENTION: THE SAME
; FILE REFERENCE: 21402-290D
; CURRENT APPLICATION NUMBER: US/10/094,466
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/288,148
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: 60/274,849
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/275,235
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/338,375
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/275,579
; PRIOR FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/335,302
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/275,601
; PRIOR FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/276,000
; PRIOR FILING DATE: 2001-03-14

```
; PRIOR APPLICATION NUMBER: 60/277,338
; PRIOR FILING DATE: 2001-03-20
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patin 2.1
; SEQ ID NO 38
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-466-38

Query Match      32.3%; Score 247.5; DB 15; Length 403;
Best Local Similarity 42.0%; Pred. No. 5.3e-17;
Matches 55; Conservative 28; Mismatches 43; Indels 5; Gaps 1;

QY 1 DLMHIVQADNPISVVEECLEAFKQVFGSLERRTAQVRYLKPQYEEGKVSAYVLRLETL 60
Db 96 EVMRLQANPNLSVADFRLAMKLVFGESSESVTAHGKFFNTLQAQGEKASLYVIRLEVQ 155
QY 61 LRAVEKRAIPRIADQVLEQVMAGATLNQMLWCRLREL-----KDQGPSPSFLLELMKV 115
Db 156 LQNAIQAGLIAEKDANRTLQQLLGAELNRDLRFRLKHLRLMYANKQERLPNFLELIKM 215
QY 116 IREEEEEEASF 126
Db 216 IREEDWDADF 226

RESULT 15
US-10-296-115-1208
; Sequence 1208, Application US/10296115
; Publication No. US20040053248A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: No. US20040053248A1e1 Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: US/10/296,115
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 1208
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-296-115-1208

Query Match      31.3%; Score 240; DB 12; Length 337;
Best Local Similarity 37.2%; Pred. No. 2.5e-16;
Matches 54; Conservative 34; Mismatches 45; Indels 12; Gaps 2;

QY 1 DLMHIVQADNPISVVEECLEAFKQVFGSLERRTAQVRYLKPQYEEGKVSAYVLRLETL 60
Db 80 EVMRVLIQATNPNSLVADFRLAMKLVFGESSESVTAHGKFFNTLQAQGEKASLYVIRLEVQ 139
QY 61 LRAVEKRAIPRIADQVLEQVMAGATLNQMLWCRLREL-----KDQGPSPSFLLELMKV 115
Db 140 LQNAIQAGITIAEKDANRTLQQLLGGELSRDLRLKDFLRMYANEQERLPNFLELIKM 199
QY 116 IREEEEEEASF-----ENESIEE 133
Db 200 VREEDWDADFIRKRPKRSSESV 224

Search completed: September 27, 2004, 17:26:04
Job time : 44.3366 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 27, 2004, 17:06:36 ; Search time 7.54739 Seconds
(without alignments)
1019.197 Million cell updates/sec

Title: US-10-037-860-9

Perfect score: 766

Sequence: 1 DLMHIVQADNPISIVECLE.....SIEEPERDGYGRWNHEGDD 149

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*

6: /cgn2_6/ptodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	323	42.2	462	4	US-09-189-527-13
2	256.5	33.5	329	4	US-09-189-527-4
3	89.5	11.7	800	4	US-09-555-790A-2
4	81.5	10.6	1898	1	US-08-056-200-94
5	81.5	10.6	1898	2	US-08-800-644-94
6	80	10.4	825	4	US-09-540-824-26
7	77	10.1	237	2	US-08-469-537A-85
8	77	10.1	370	3	US-08-857-076-107
9	77	10.1	661	4	US-09-107-532A-3677
10	77	10.1	1367	2	US-08-249-687C-2
11	77	10.1	1367	2	US-08-625-819-2
12	77	10.1	1367	3	US-08-746-559A-2
13	77	10.1	1367	4	US-08-864-641B-18
14	77	10.1	1367	4	US-09-343-551-2
15	76	9.9	754	4	US-09-585-173B-51
16	75.5	9.9	591	4	US-09-198-452A-806
17	75	9.8	571	4	US-09-489-039A-10805
18	75	9.8	704	4	US-09-370-838-191
19	75	9.8	1197	4	US-09-618-425-2
20	74	9.7	217	1	US-08-185-424B-2
21	74	9.7	322	4	US-09-540-236-2700
22	74	9.7	420	3	US-09-329-418-8
23	74	9.7	420	3	US-09-531-914-8
24	74	9.7	497	4	US-09-345-473E-8
25	74	9.7	518	3	US-09-329-418-3
26	74	9.7	518	3	US-09-329-418-4
27	74	9.7	518	3	US-09-329-418-5

28 74 9.7 518 3 US-09-329-418-9 Sequence 9, Appli
29 74 9.7 518 3 US-09-531-914-3 Sequence 3, Appli
30 74 9.7 518 3 US-09-531-914-4 Sequence 4, Appli
31 74 9.7 518 3 US-09-531-914-5 Sequence 5, Appli
32 74 9.7 518 3 US-09-531-914-9 Sequence 9, Appli
33 73.5 9.6 336 4 US-09-252-991A-24656 Sequence 24656, A
34 73.5 9.6 783 6 5231168-2 Patent No. 5231168
35 73.5 9.6 1208 4 US-09-134-000C-5756 Sequence 5756, Ap
36 73 9.5 355 4 US-09-134-001C-3622 Sequence 3622, Ap
37 73 9.5 1588 5 PCT-US93-07261-11 Sequence 11, Appl
38 73 9.5 1663 5 PCT-US93-07261-16 Sequence 16, Appl
39 73 9.5 1985 4 US-09-495-714C-6 Sequence 6, Appli
40 72.5 9.5 258 4 US-09-489-039A-12052 Sequence 12052, A
41 72.5 9.5 651 3 US-08-650-766-6 Sequence 5, Appli
42 72.5 9.5 651 3 US-08-922-635-5 Sequence 5, Appli
43 72.5 9.5 651 4 US-09-389-487-6 Sequence 6, Appli
44 72.5 9.5 1070 3 US-08-922-635-22 Sequence 22, Appli
45 72.5 9.5 1504 4 US-09-364-206-2 Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-189-527-13
; Sequence 13, Application US/09189527A

; Patent No. 6387639

; GENERAL INFORMATION:

; APPLICANT: Jerome B. Posner

; APPLICANT: Joseph O. Dalmau

; APPLICANT: Myrna R. Rosenfeld

; TITLE OF INVENTION: Ma Family Polypeptides and Anti-Ma

; TITLE OF INVENTION: Antibodies

; FILE REFERENCE: SLK98-01

; CURRENT APPLICATION NUMBER: US/09/189,527A

; CURRENT FILING DATE: 1998-11-10

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 13

; LENGTH: 462

; TYPE: PRT

; ORGANISM: homo sapiens

US-09-189-527-13

Query Match 42.2%; Score 323; DB 4; Length 462;

Best Local Similarity 52.7%; Pred. No. 2.9e-29;

Matches 68; Conservative 26; Mismatches 33; Indels 2; Gaps 1;

Qy 6 VQADNPISIVECLEAPKQVFGSLESRRTAQVRVLYKPYQEGEKVSAYVLRLETLRRAY 65

Db 214 LRASNASITVEECLEALQQVFGPVESHKIAQVKLCAYQEGEKVSFVLRLEPLQRAV 273

Qy 66 EKRAIPRIADQVRLEOVNMGATINQMLCRLRELKDGPPPSFLELMKVIREEESEAS 125

Db 274 ENNVSRNVNQTRLKRVLSGATLPDKLRDKLKLKQRRKPPGFLALVKLLREEEWEAT 333

Qy 126 F--ENESIE 132

Db 334 LGPDRESLE 342

RESULT 2

US-09-189-527-4

; Sequence 4, Application US/09189527A

; Patent No. 6387639

; GENERAL INFORMATION:

; APPLICANT: Jerome B. Posner

; APPLICANT: Joseph O. Dalmau

; APPLICANT: Myrna R. Rosenfeld

; TITLE OF INVENTION: Ma Family Polypeptides and Anti-Ma

; TITLE OF INVENTION: Antibodies

; FILE REFERENCE: SLK98-01

; CURRENT APPLICATION NUMBER: US/09/189,527A

```
; CURRENT FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 329
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-189-527-4

Query Match      33.5%; Score 256.5; DB 4; Length 329;
Best Local Similarity 46.6%; Pred. No. 1.1e-21;
Matches 54; Conservative 25; Mismatches 30; Indels 7; Gaps 2;

QY 1 DLMHIVQADNPSISVECLAEKQVFGSLERTRTAQVRLKPYQEGEKVSAYVLRLEWL 60
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 212 DVIRILKNNPAITTAELKALEQVFGSVESRDQIKFLNTYQNPGEKLSAYVIRLEP 271
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 61 LRAVRAKRAIPRIADQVRLQVWAGA-----TLNQLWLCRLRELKDQGGPPPSFLEL 112
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 272 LQKVEKGAIDKNVQARLEQVIAGANHSGAIRQLWL---TGAGEGPGKPLSV 324
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 3
US-09-555-790A-2
; Sequence 2, Application US/09555790A
; Patent No. 6555652
; GENERAL INFORMATION:
; APPLICANT: ITOH, Kyogo et al.
; TITLE OF INVENTION: TUMOR ANTIGEN PEPTIDE DERIVATIVES
; FILE REFERENCE: 0020-4716P
; CURRENT APPLICATION NUMBER: US/09/555,790A
; CURRENT FILING DATE: 2000-07-12
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 2
; LENGTH: 800
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-555-790A-2

Query Match      11.7%; Score 89.5; DB 4; Length 800;
Best Local Similarity 25.3%; Pred. No. 0.13;
Matches 37; Conservative 25; Mismatches 51; Indels 33; Gaps 6;

QY 5 IVQADNPSISVECLAEKQVFGSLERTRTAQVRLKPYQEGEKVSAYVLRLEL 64
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 489 VLEDEAELELQKLE-----KGRRLRLQLQLRDSGKQVIVKLESQRGW 539
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 65 VKRAIPRIADQVRLQVWAGATLNQMLWCR-LRELKDQGGPPPSFLELMKVIREEEEE 123
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 540 EEDE-----DPERKGAIVFNATSE---FCRTLGLPIPTYG-----LAGNREEQEEL 581
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 124 ASPENESIEEPEDRGYGRNHEGDD 149
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 582 MDEFD-----EERSANGGESDGE 602
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 4
US-08-056-200-94
; Sequence 94, Application US/08056200
; Patent No. 5616500
; GENERAL INFORMATION:
; APPLICANT: Steinert, Peter M.
; APPLICANT: Lee, Seung-Chul
; APPLICANT: Kim, In-Gyu
; APPLICANT: Chung, Soo-Il
; APPLICANT: Park, Sang-Chul
; TITLE OF INVENTION: Trichohyalin and Transglutaminase-3 and
; TITLE OF INVENTION: Methods of Using Same
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/800,644
; FILING DATE: 14-FEB-1997
```

```
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA: US 08/056,200
; APPLICATION NUMBER: 30-APR-1993
; FILING DATE: 30-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Fredrick, Michael F.
; REGISTRATION NUMBER: 36,799
; REFERENCE/DOCKET NUMBER: NIH054.001A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (714) 760-0404
; TELEFAX: (714) 760-9502
; INFORMATION FOR SEQ ID NO: 94:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1898 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-800-644-94

Query Match      10.6%; Score 81.5; DB 2; Length 1898;
Best Local Similarity 28.9%; Pred. No. 3.9;
Matches 33; Conservative 16; Mismatches 45; Indels 25; Gaps 4;

Qy 20 EAFKQVGSLSRRTAQR---YLPKQEAGEKVSAYVLRLETLRLRAVEKRAIPRIAD 76
Db 563 ERLQQLKREKREKLEQERREQLKREQE-----REDQLKREERQQLKREQ 613

Qy 77 QVRLEQVMAGATNQLMWCRLEKLDQGPFPFLMKVIREEEERASFNESIEPPE 136
Db 614 EERLEQLKREVERL---EQEERRDE-----RLKREPEERHRLKSERQEE 660

Qy 137 R 137
Db 561 R 661

RESULT 6
US-09-540-824-26
; Sequence 26, Application US/09540824
; GENERAL INFORMATION:
; APPLICANT: Thiele, Dennis
; TITLE OF INVENTION: Liu, Phillip
; FILE REFERENCE: UM-04266
; CURRENT APPLICATION NUMBER: US/09/540,824
; CURRENT FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 26
; LENGTH: 825
; TYPE: PRP
; ORGANISM: Schizosaccharomyces pombe
US-09-540-824-26

Query Match      10.4%; Score 80; DB 4; Length 825;
Best Local Similarity 25.8%; Pred. No. 1.8;
Matches 33; Conservative 25; Mismatches 32; Indels 38; Gaps 8;

Qy 30 ESRRTAQRYLKPQEAGEKVSAYVLRLETLRLRAV-EKRAIP-RR1ADQVRLEQVMAGA 87
Db 253 KASRKSGIKTQRP1ISGD-----ARYDSFVREWFDKRAHPTERTKTEELAQIAD- 305

Qy 88 TLNQLMWCRLEKLDQGPFPFLMKVIR-EEEEERASFNESIEPPEERD---GYGR- 142
Db 306 -----RLRELEDO-----RISRMHYQEDSASEAGSIEDEQATDNVFGFGK 347

Qy 143 -----WN 144
Db 348 QENEENWN 355

RESULT 7
```

```
US-08-469-537A-85
; Sequence 85, Application US/08469537A
; Patent No. 5843749
; GENERAL INFORMATION:
; APPLICANT: Maisonnier, et al.
; TITLE OF INVENTION: EHK AND ROR TYROSINE
; TITLE OF INVENTION: KINASES
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Regeneron Pharmaceuticals, Inc.
; STREET: 777 Old Saw Mill River Road
; CITY: Tarrytown
; STATE: NY
; COUNTRY: U.S.A.
; ZIP: 10591
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,537A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/406,247
; FILING DATE: 17-MAR-1995
; APPLICATION NUMBER: USSN 08/144,992
; FILING DATE: 28-OCT-1993
; APPLICATION NUMBER: USSN 07/736,559
; FILING DATE: 26-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Kempner, Ph.D., Gail M
; REGISTRATION NUMBER: 32,143
; REFERENCE/DOCKET NUMBER: REG 070C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 914-345-7400
; TELEFAX: 914-345-7721
; TELEX:
; INFORMATION FOR SEQ ID NO: 85:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 237 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-469-537A-85

Query Match      10.1%; Score 77; DB 2; Length 237;
Best Local Similarity 34.2%; Pred. No. 0.66;
Matches 27; Conservative 12; Mismatches 26; Indels 14; Gaps 4;

Qy 74 IADQVRLEQVMAGATLNQ-----MLWCRLRELKLDQGP--PSFLELMKVIREEE---E 122
Db 118 LSNEQVLRFVMEGGLDKPCDMLFELMRMCQYNPKRPSFLELISSIKKEEMEGFR 177

Qy 123 EASF---ENSEIEPBERD 138
Db 178 EVSFYTYSEENKLPPEELD 196

RESULT 8
US-08-857-076-107
; Sequence 107, Application US/08857076C
; Patent No. 6225120
; GENERAL INFORMATION:
; APPLICANT: Ruvkun, Gary
; APPLICANT: Kimura, Koutarou
; APPLICANT: Patterson, Garth
; APPLICANT: Ogg, Scott
; APPLICANT: Paradis, Suzanne
; APPLICANT: Tissenbaum, Heidi
; APPLICANT: Morris, Jason
```

```
; APPLICANT: Kowek, Allison
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
; FILE REFERENCE: 00786/351001
; CURRENT APPLICATION NUMBER: US/08/857,076C
; CURRENT FILING DATE: 1997-05-15
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 107
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-857-076-107

Query Match      10.1%; Score 77; DB 3; Length 370;
Best Local Similarity 34.2%; Pred. No. 1.3;
Matches 27; Conservative 12; Mismatches 26; Indels 14; Gaps 4;

QY 74 IADQVRLEQVMAGATLQ-----MLWCRLRELKQGGP--PSFLELMKVIREEEE---E 122
Db 278 LSNEQVLRFVMEGLLKDPCDMLFELMRMCWQYNPKMRPSFLELSISKEEMEPGPR 337

QY 123 EASF---ENESIEEPERD 138
Db 338 EVSFYVSEENKLPPEELD 356

RESULT 9
US-09-107-532A-3677
; Sequence 3677, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESS: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3677:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 661 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
```

```
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...661
; SEQUENCE DESCRIPTION: SEQ ID NO: 3677:
US-09-107-532A-3677

Query Match      10.1%; Score 77; DB 4; Length 661;
Best Local Similarity 23.5%; Pred. No. 2.9;
Matches 31; Conservative 27; Mismatches 54; Indels 20; Gaps 4;

QY 9 DNPSSIVEECLAEAFKQVFGSLERSTAQVRYLKPYQEEGKYSAYVLRLETLRLRAVEKR 68
Db 512 DVPEVSLVAILDADKE--GFLRSERSLVQTCIGRAARNBEGKVIYADKVTDSMRAMDST 569

QY 69 AIPRIADQVRLEQVMAGATLQMLWCRLRELKQGGPPSFLELMKVIREEEEEASPEN 128
Db 570 SRRRTIQKYNHEHGIIVKTI-----IKEIRD-----LISITKESEDDTKEAVQ 613

QY 129 ESIEE--PEERD 138
Db 614 VSYEEMTKEEKD 625

RESULT 10
US-08-249-687C-2
; Sequence 2, Application US/08249687C
; Patent No. 5942412
; GENERAL INFORMATION:
; APPLICANT: PRAGER, DIANE
; APPLICANT: MELMED, SHLOMO
; TITLE OF INVENTION: POLYNUCLEIC ACID ENCODING
; TITLE OF INVENTION: VARIANT INSULIN-LIKE GROWTH FACTOR I RECEPTOR BETA
; TITLE OF INVENTION: SUBUNIT & RECEPTOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder & Poplawski
; STREET: 444 South Flower St., 19th Floor
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/249,687C
; FILING DATE: 26-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/044,540
; FILING DATE: 06-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Basile, Lena
; REGISTRATION NUMBER: P-44,026
; REFERENCE/DOCKET NUMBER: P07 32349
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 213-622-7700
; TELEFAX: 213-489-4210
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1367 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-249-687C-2

Query Match      10.1%; Score 77; DB 2; Length 1367;
Best Local Similarity 34.2%; Pred. No. 8.1;
Matches 27; Conservative 12; Mismatches 26; Indels 14; Gaps 4;
```



```

QY 74 IAOVRLQVWAGATLNQ-----MLWCRLRLKQOGPP--PSFLEMKVIREEE---E 122
      :::: ||| ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1216 LSNEQVLRVFWGGGLDKDPCNDPMFELRMCMQYNPKRPSFLEIISIKKEMEGFR 1275
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 123 EASF--ENESIEEPEERD 138
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1276 EVSFYFYEENKLEPEERD 1294
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

```

```

RESULT 11
US-08-625-819-2
; Sequence 2, Application US/08625819
; Patent No. 5958872
; GENERAL INFORMATION:
; APPLICANT: O'CONNOR, Rosemary; and
; APPLICANT: BASERGA, Renato L.
; TITLE OF INVENTION: ACTIVE SURVIVAL DOMAINS OF IGF-IR
; TITLE OF INVENTION: AND METHODS OF USE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
;

```

Query Match	10.1%	Score 77;	DB 2;	Length 1367;
Best Local Similarity	34.2%;	Pred. No. 8.1;		
Matches	27;	Conservative 12;	Mismatches 26;	Indels 14;
Gaps	4;			

QY	74	IADQVRLVQMGATLNQ-----MLWCRLRLKQDQPP--PQFLELMKVIREEE---E 122
Db	1216	LSNEQVLRFVMEGLDKDPCNDMLFELMRMCWQYNPKRPFLEIISIIKKEMEGFR 1275
QY	123	EASF---ENESTEEPERD 138
Db	1276	EVSFYFSEENKLPPEPELD 1294

RESULT 12
US-08-746-559A-2
; Sequence 2, Application US/08746559A
; Patent No. 6084085
; GENERAL INFORMATION:
; APPLICANT: Renato Baserga
; APPLICANT: Mariana Resnicoff
; APPLICANT: Consuelo D'Ambrosio
; APPLICANT: Andre Ferber
; TITLE OF INVENTION: Method of Inc

NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 6084085
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: WORDPERFECT 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/746,559A
FILING DATE: 13-NOV-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/006,699
FILING DATE: 14-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: Paul K. Legaard
REGISTRATION NUMBER: 38,534
REFERENCE/DOCKET NUMBER: TJU-2063
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1367 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-746-559A-2

Query Match 10.1%; Score 77; DB 3; Length 1367;
Best Local Similarity 34.2%; Pred. No. 8.1;
Matches 27; Conservative 12; Mismatches 26; Indels

```

Qy 74 IADQVLEQVMGATUNQ-----MLWCKRLRELKDGPP--PSFLEMLKVIREEE--E 122
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1216 LSNEQVLRFMVEGLLDKPDNCDFMLFELMRWCQYNKMRPSFELEIISIKEMERGFR 1275
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 123 EASF---ENESIEEPEERD 138
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1276 EVSFYIYSEENKLPPEELD 1294
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 13
US-08-864-641B-18
; Sequence 18, Application US/08864641B
; Patent No. 6312684
; GENERAL INFORMATION:
; APPLICANT: Baserga, Renato
; APPLICANT: Abraham, David
; APPLICANT: Resnicoff, Mariana
; TITLE OF INVENTION: Method Of Inducing Resistance To Tumor Growth
; FILE REFERENCE: TJU2137
; CURRENT APPLICATION NUMBER: US/08/864,641B
; CURRENT FILING DATE: 1997-05-29
; PRIOR APPLICATION NUMBER: 08/340,732
; PRIOR FILING DATE: 1994-11-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 1367
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: No. 6312684el Sequence
US-08-864-641B-18

```

Query Match

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 27, 2004, 17:02:20 ; Search time 25.0768 Seconds
(without alignment)
1678.826 Million cell updates/sec

Title: US-10-037-860-9

Perfect score: 766

Sequence: 1 DLHIVQADNPISVEGLE.....SIEPPEERDGYGRNHEGDD 149

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_29Jan04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	766	100.0	149	3 AAB12527	Aab12527 Human Ma3
2	758	99.0	364	7 ADC08977	Adc08977 Oncogene
3	755	98.6	283	3 AAB12528	Aab12528 Human Ma4
4	323	42.2	455	5 ABB05727	Abb05727 Human sig
5	323	42.2	463	3 AAB12529	Aab12529 Human Ma5
6	323	42.2	463	3 AAB42315	Aab42315 Human ORF
7	295.5	38.6	353	4 AAB74701	Aab74701 Human mem
8	287	37.5	452	6 AAO16179	Aao16179 Human pro
9	277	36.2	399	6 ADA54410	Ada54410 Human pro
10	277	36.2	399	6 ABG99947	Abg99947 Human nov
11	271.5	35.4	280	4 AAE01336	Aae01336 Human gen
12	271.5	35.4	351	4 AAE01340	Aae01340 Human gen
13	271.5	35.4	351	4 AAB74695	Aab74695 Human mem
14	271.5	35.4	351	4 AAU08664	Aau08664 Human NOV
15	266.5	34.8	237	4 AAB94854	Aab94854 Human pro
16	253.5	33.1	329	3 AAB12525	Aab12525 Human Ma1
17	247.5	32.3	403	5 ABG97495	Abg97495 Human NOV
18	247.5	32.3	403	5 AAMS1624	Aams1624 KIAA0883
19	247.5	32.3	403	6 ABO14772	Abol14772 Novel hum
20	246.5	32.2	403	6 ABO14773	Abol14773 Novel hum
21	240	31.3	337	4 AAM25693	Aam25693 Human pro
22	237	30.9	402	4 AAB60478	Aab60478 Human cel
23	225	29.4	312	3 AAB43023	Aab43023 Human ORF
24	180	23.5	80	4 ABB15036	Abb15036 Human ner
25	125	16.3	110	4 AAO01787	Aao01787 Human pol

26 99.5 13.0 1270 6 ABO14659 Novel hum
27 98.5 12.9 525 5 ABB57353 Mouse isc
28 97.5 12.7 218 6 ADA55313 Human pro
29 97.5 12.7 543 4 AAM93466 Human pol
30 97.5 12.7 543 7 ADE57654 Human pro
31 96.5 12.6 541 7 ADE57652 Rat Prote
32 96.5 12.6 542 6 ABO4812 Rat NF-L
33 95.5 12.5 411 2 AAM24229 Human mel
34 94.5 12.3 475 2 AAW24228 Human mel
35 94.5 12.3 525 2 AAW24230 Human mel
36 90 11.7 592 5 ABG30418 Human tes
37 90 11.7 592 6 ABU00143 Human nov
38 89.5 11.7 563 2 AAW64228 Human sec
39 89.5 11.7 563 4 AAB90738 Human C22
40 89.5 11.7 760 5 ABB90912 Herbicida
41 89.5 11.7 760 7 ADB95024 A. thalia
42 89.5 11.7 800 2 AAW44003 A tumour
43 89.5 11.7 800 2 AAY27254 Human tum
44 89.5 11.7 800 4 AAB60503 Human cel
45 89.5 11.7 800 6 ABO53005 Human spl

ALIGNMENTS

RESULT 1
AAB12527

ID AAB12527 standard; protein; 149 AA.

XX AAB12527;

DT 02-NOV-2000 (first entry)

XX Human Ma3 protein SEQ ID NO:9.

XX Mal; Ma2; Ma3; Ma4; Ma5; paraneoplastic limbic encephalitis; diagnosis;
XX paraneoplastic syndrome; neoplasm; colon cancer; brainstem encephalitis;
XX breast cancer; parotid gland cancer; lung cancer; testicular cancer;
XX germ-cell tumour.

XX Homo sapiens.

XX JP2000146982-A.

XX 26-MAY-2000.

XX 10-NOV-1999; 99JP-00320171.

XX 10-NOV-1998; 98US-00189527.

XX (SLOK) SLOAN KETTERING INST CANCER RES.

XX WPI; 2000-468119/41.

XX N-PSDB; AAA60835.

XX Diagnosing paraneoplastic syndrome e.g. paraneoplastic limbic
XX encephalitis or neoplasm e.g. colon cancer comprising assessing a test
XX sample for the presence or absence of antibodies to a Ma family
XX polypeptide.

XX Claim 48; Fig 6; 27pp; Japanese.

XX The present invention describes a method for diagnosing a paraneoplastic
XX syndrome or neoplasm. The method comprises assessing a test sample for
XX the presence or absence of antibodies to a Ma family polypeptide (I). The
XX method is used to diagnose a paraneoplastic syndrome especially
XX paraneoplastic limbic encephalitis and/or brainstem encephalitis or
XX neoplasm in an individual. The method diagnoses the neoplasm by assessing
XX antibodies to (I) preferably Ma1, which is indicative presence of breast
XX cancer, colon cancer, parotid gland cancer, lung cancer, testicular
XX cancer and germ-cell tumours or Ma2, which is indicative of testicular
XX cancer, germ-cell tumour, and lung cancer. The present sequence is the
XX Ma3 protein as given in the present invention

XX SQ Sequence 149 AA;

Query Match 100.0%; Score 766; DB 3; Length 149;
 Best Local Similarity 100.0%; Pred. No. 1.8e-75;
 Matches 149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLHIVQADNPISVVEECLEAFKQVFGSLESRRTAQVRLKPYQEGEKVSAYVLRLETL 60
 DB 1 DLHIVQADNPISVVEECLEAFKQVFGSLESRRTAQVRLKPYQEGEKVSAYVLRLETL 60

QY 61 LRRAVEKAIARRIADQVRLQVWAGATLNQMLWCLRELKDKQPPPSFLELMKVIREEE 120
 DB 61 LRRAVEKAIARRIADQVRLQVWAGATLNQMLWCLRELKDKQPPPSFLELMKVIREEE 120

QY 121 EEEASFENESIEEPEERDGYGRWNHEGDD 149
 DB 121 EEEASFENESIEEPEERDGYGRWNHEGDD 149

RESULT 2
 ADC08977
 ID ADC08977 standard; protein; 364 AA.
 XX
 AC ADC08977;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DE Onconeural antigen Ma2 protein.
 XX
 KW Human; Ma2; onconeural; antigen; Alzheimer's disease;
 KW neurodegenerative disease; diagnosis; neuroprotective; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO2003073104-A2.
 XX
 PD 04-SEP-2003.
 XX
 PF 26-FEB-2003; 2003WO-BP001946.
 XX
 PR 26-FEB-2002; 2002EP-00004177.
 PR 26-FEB-2002; 2002US-0359307P.
 XX
 PA (EVOT-) EVOTEC NEUROSCIENCES GMBH.
 XX
 PI Hipfel R, Von Der Kammer H, Pohlner J;
 XX
 XX WPI; 2003-721818/68.
 DR GENBANK; O94959, KIA00883.
 XX
 XX Diagnosing or prognosticating, or determining increased risk of
 PT developing a neurodegenerative disease by determining level or activity
 PT of a transcription or translation product of a gene coding for Ma
 PT onconeural antigen.
 XX
 PS Disclosure; Fig 9; 51pp; English.
 XX
 XX The present sequence is the protein sequence of human onconeural
 CC antigen Ma2. The invention discloses the detection and differential
 CC expression and regulation of the Ma2 gene in specific brain regions of AD
 CC patients. The Ma2 gene and its transcription and/or translation products
 CC may have a causative role in the regional selective neuronal degeneration
 CC typically observed in AD, or may confer a neuroprotective function to the
 CC remaining nerve cells. Methods are claimed for diagnosing or
 CC prognosticating a neurodegenerative disease, for monitoring the
 CC progression of a neurodegenerative disease, and for evaluating treatment
 CC of a neurodegenerative disease, especially AD, in a subject by
 CC determining the level and/or activity of a transcription or translation
 CC product of an Ma onconeural antigen gene, especially Ma2. Also claimed
 CC are: a method for treating or preventing AD and related neurodegenerative
 CC disorders using the Ma2 gene or its transcription or translation product;
 CC a method of screening for modulating agents of neurodegenerative diseases

CC ; and a recombinant non-human animal comprising an Ma2 gene sequence,
 CC which is useful for screening, testing and validating candidate
 CC diagnostic and therapeutic agents.
 XX
 SQ Sequence 364 AA;

Query Match 99.0%; Score 758; DB 7; Length 364;
 Best Local Similarity 99.3%; Pred. No. 4.4e-74;
 Matches 148; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DLHIVQADNPISVVEECLEAFKQVFGSLESRRTAQVRLKPYQEGEKVSAYVLRLETL 60
 DB 216 DLHIVQADNPISVVEECLEAFKQVFGSLESRRTAQVRLKPYQEGEKVSAYVLRLETL 275

QY 61 LRRAVEKAIARRIADQVRLQVWAGATLNQMLWCLRELKDKQPPPSFLELMKVIREEE 120
 DB 276 LRRAVEKAIARRIADQVRLQVWAGATLNQMLWCLRELKDKQPPPSFLELMKVIREEE 335

QY 121 EEEASFENESIEEPEERDGYGRWNHEGDD 149
 DB 336 EEEASFENESIEEPEERDGYGRWNHEGDD 364

RESULT 3
 AAB12528
 ID AAB12528 standard; protein; 283 AA.
 XX
 AC AAB12528;
 XX
 DT 02-NOV-2000 (first entry)
 XX
 DE Human Ma4 protein SEQ ID NO:11.
 XX
 KW Ma1; Ma2; Ma3; Ma4; Ma5; paraneoplastic limbic encephalitis; diagnosis;
 KW paraneoplastic syndrome; neoplasm; colon cancer; brainstem encephalitis;
 KW breast cancer; parotid gland cancer; lung cancer; testicular cancer;
 KW germ-cell tumour.
 XX
 OS Homo sapiens.
 XX
 PN JP2000146982-A.
 XX
 PD 26-MAY-2000.
 XX
 PF 10-NOV-1999; 99JP-00320171.
 XX
 PR 10-NOV-1998; 98US-00189527.
 XX
 PA (SLOK) SLOAN KETTERING INST CANCER RES.
 XX
 XX WPI; 2000-468119/41.
 DR N-PSDB; AAM60836.
 XX
 XX Diagnosing paraneoplastic syndrome e.g. paraneoplastic limbic
 PT encephalitis or neoplasm e.g. colon cancer comprising assessing a test
 PT sample for the presence or absence of antibodies to a Ma family
 PT polypeptide.
 XX
 PS Claim 48; Fig 7-8; 27pp; Japanese.
 XX
 XX The present invention describes a method for diagnosing a paraneoplastic
 CC syndrome or neoplasm. The method comprises assessing a test sample for
 CC the presence or absence of antibodies to a Ma family polypeptide (I). The
 CC method is used to diagnose a paraneoplastic syndrome especially
 CC paraneoplastic limbic encephalitis and/or brainstem encephalitis or
 CC neoplasm in an individual. The method diagnoses the neoplasm by assessing
 CC antibodies to (I) preferably Ma1, which is indicative presence of breast
 CC cancer, colon cancer, parotid gland cancer, lung cancer, testicular
 CC cancer and germ-cell tumours or Ma2, which is indicative of testicular
 CC cancer, germ-cell tumour, and lung cancer. The present sequence is the
 CC Ma4 protein as given in the present invention
 XX
 SQ Sequence 283 AA;

```

Query Match      98.6%; Score 755; DB 3; Length 283;
Best Local Similarity 98.7%; Pred. No. 6.7e-74;
Matches 147; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DLHIVQADNPISVEECLEAFKQVFGSLESRRTAQVRLKPYQEGEKVSAYVLRLETL 60
Db 135 DLHIVQADNPISVEECLEAFKQVFGSLESRRTAQVRLKPYQEGEKVSAYVLRLETL 194

QY 61 LRAVEKRAIPRIADQVRLQVLMAGATLNQMLWCRRLRELKQDQPPPSFLELMKVIREEE 120
Db 195 LRAVEKRAIPRIADQVRLQVLMAGATLNQMLWCRRLRELKQDQPPPSFLELMKVIREEE 254

QY 121 BEEASFENESIEPERDQYGRWNHEGDD 149
Db 255 BEEASFENESIEPERDQYGRWNHEGDD 283

RESULT 4
ABB05727
ID ABB05727 standard; protein; 455 AA.
XX
AC ABB05727;
XX
DT 30-APR-2002 (first entry)
XX
DE Human signal transduction protein clone tes3_5k22.
XX
KW Human; foetal brain; foetal kidney; melanoma; testis; amygdala;
KW gene therapy.
XX
OS Homo sapiens.
XX
WO200198454-A2.
XX
27-DEC-2001.
XX
25-APR-2001; 2001WO-IB002050.
XX
25-APR-2000; 2000US-0199380P.
XX
(GEHU-) GERMAN HUMAN GENOME PROJECT.
XX
Wiemann S;
XX
WPI; 2002-055860/07.
XX
N-PSDB; ABA93764.
XX
Human cDNA sequences and clones derived from human fetal brain, fetal
screening and therapy.
XX
Claim 1; Page 382; 61pp; English.
XX
The present invention describes assemblages and computer readable media
comprising novel human cDNA sequences and clones derived from human
foetal brain, foetal kidney, melanoma, testis and amygdala cDNA
libraries. ABA93702 to ABA93766 represent human cDNA sequences from the
present invention which encode the proteins given in ABB05662 to
ABB05729. The human cDNA sequences and clones can be used in gene
therapy. The clones may be used in a variety of applications, for example
they may be used in profiling assays, for providing large arrays of human
genetic material for implementing large-scale screening strategies and
for treating diseases via gene therapy procedures
XX
Sequence 455 AA;

Query Match      42.2%; Score 323; DB 5; Length 455;
Best Local Similarity 52.7%; Pred. No. 2.1e-26;
Matches 68; Conservative 26; Mismatches 33; Indels 2; Gaps 1;

QY 6 VQADNPISVEECLEAFKQVFGSLESRRTAQVRLKPYQEGEKVSAYVLRLETLRRAY 65
Db 220 LRASNASITVEECLEAALQVFGFVESHKTAQVRLKPYQEGEKVSAYVLRLETLRRAY 279

QY 66 EKRAIPRIADQVRLQVLMAGATLNQMLWCRRLRELKQDQPPPSFLELMKVIREEEAS 125
Db 280 ENNVSRNVNQTRLKRVLSGATLPDKLRDKLKMQRKPPGFLALVKLLREEEWEAT 339

Query Match      42.2%; Score 323; DB 3; Length 463;
Best Local Similarity 52.7%; Pred. No. 2.1e-26;
Matches 68; Conservative 26; Mismatches 33; Indels 2; Gaps 1;

QY 6 VQADNPISVEECLEAFKQVFGSLESRRTAQVRLKPYQEGEKVSAYVLRLETLRRAY 65
Db 220 LRASNASITVEECLEAALQVFGFVESHKTAQVRLKPYQEGEKVSAYVLRLETLRRAY 279

QY 66 EKRAIPRIADQVRLQVLMAGATLNQMLWCRRLRELKQDQPPPSFLELMKVIREEEAS 125
Db 280 ENNVSRNVNQTRLKRVLSGATLPDKLRDKLKMQRKPPGFLALVKLLREEEWEAT 339

```

CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
 CC cartilage damage, nocturnal haemoglobinuria, anti-inflammatory disease; to
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive
 XX
 XX Sequence 463 AA;
 SO

Query Match	42.2%;	Score 323;	DB 3;	Length 463;
Best Local Similarity	52.7%;	Pred. No. 2.1e-26;		
Matches 68;	Conservative	26;	Mismatches 33;	Indels 2;
				Gaps 1

[illegible]

66 EKRAIPRIADQVRLSEQVMAGATINQMLWCRLRELKDQGGPPPSFLFELMKVIREEEEEAS 125

Qy 126 F--ENESIE 132
: |||

DD 340 LGFKRESLE 340
PBCNLT 7

AA74701	ID	standard; protein; 353 AA.
XX	AA74701	

XX	12-JUN-2001	(first entry)
XX		
XX		
XX		

XX Human; membrane associated protein; MEMAP; diagnosis; cytostatic;
KW anti-inflammatory; anticonvulsant; immunosuppressive; antiarrhetic;
KW Human; membrane associated protein; MEMAP; diagnosis; cytostatic;
XX

KW antiatherosclerotic; gene therapy; cell proliferative disorder;
KW autoimmune disorder; inflammatory disorder; neurological disorder;
KW gastrointestinal disorder; cancer; inflammation; atherosclerosis;
KW gastrointestinal disorder; cancer; inflammation; atherosclerosis;

XX
OS
XX
Homo sapiens.
XX

FR
XX
XX
PD
XX
XX

WOLVERINE INC.
22-FEB-2001.

PF	14-AUG-2000; 2000WO-US02Z315.
XX	
PR	17-AUG-1999; 99US-0149641P.
CR	06 NOV-1998; 99US-0164203D.

XX
PA
XX
(INCY-) INCYTE GENOMICS INC.

FI Dear Sir,
PI I am in Lu
PI Baughn MR, Lu DAM, Patterson C;
XX
DR WPI: 2001-168860/17.

DR N-PSDE; AAF81/47.
XX Isolated polypeptide with a human membrane associated protein sequence identical to that of the AAF81/47 polypeptide.
PT Useful for the diagnosis, prevention and treatment of cell proliferative disorders.

PT autoimmune/inflammatory, neurological and gastrointestinal disorders.
XX
PS Claim 1; Page 119-120; 173pp; English.

CC AAF81741 to AAF81777 encode the human membrane associated proteins
CC (MEMAP) given in AAB74695 to AAB74731. MEMAPs have cytosolic,
CC antiinflammatory, anticonvulsant, immunosuppressive, antiarrheic and

CC MEMAPs and agonist of MEMAPs can be used to treat a disease or condition
CC associated with decreased expression of functional MEMAP and antagonists
CC of MEMAP are used to treat a disease or condition associated with
CC increased expression of functional MEMAP and agonists of MEMAP
CC can be used to treat a disease or condition associated with decreased
CC expression of functional MEMAP and antagonists of MEMAP

QY 126 F--ENSIE 132
: ||:
Db 340 LGPRESLE 348

RESULT 6
AAB42315
ID AAB42315 standard; protein; 463 AA.

AC
XX
DT
08-FEB-2001
08-FEB-2001 (first entry)
AAB42315;
AC
XX
DT

Human ORFX ORF2079 polypeptide sequence SEQ ID NO:4158:
Human: open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW

KW vulnereary; antiposioratic; antiparkinsonian; humnropic; neuroprotective;
 KW anticovulsant; osteopathic; antiarthritic; immunosuppressant; cardiac;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW humectant; dermatological; immunosuppressive; antiinflammatory;
 KW

KW antitubercular; antifungal; antirheumatic; antithyroid;
antiviral; antibacterial; cancer; proliferative disorder; hypertension;
antianemic; gene therapy; cancer; osteoarthritis, graft vs. host disease;
KW neurodegenerative disorder; osteoarthritis, graft vs. host disease;
KW neurodegenerative disorder; osteoarthritis, graft vs. host disease;

KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;

bone damage; cartilage damage; anullinflammatory disease; coagulation;
 thrombosis; contraceptive.

XX
PN
XX
WO200058473-A2.
XX
XX

31-MAR-2000; 2000WO-US008621.

2K 31-MAR-1999; 99US-0126076.
 PR 02-APR-1999; 99US-0127636P.
 PR 05-APR-1999; 99US-0127728P.
 PR 30-MAR-2000; 2000US-00540763.

XX
PA
XX
(CURA-) CURAGEN CORP.
XX
XX

XX WPI; 2000-602362/57.
DR N-PSDB; AAC76524.
DR

xx Novel nucleic acids and peptides derived from open reading frame X,
 pt useful for treating e.g. cancers, proliferative disorders,
 pt neurodegenerative disorders and cardiovascular disease.

Claim 11; Page 3345-3347; 5507pp; English.

CC which represent the human OREF open reading frames 1 to 3161. The OREF
CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
CC antiproliferative; antiparkinsonian; neurotropic; neuroprotective; osteopathic;
CC

[illegible]

CC sequences can be used for determining the presence or of preposition
CC to, or preventing or treating pathological conditions associated with an
CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
CC existing in *in vivo* therapy vectors. The proteins and nucleic acids may be

CC used to treat cancers, proliferative disorders, neurodegenerative
CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
CC

0
1
2
3
4
5
6
7
8
9
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z

CC overexpression of functional MEMAP. These disorders include cell
 CC proliferative, autoimmune/inflammatory, neurological and gastrointestinal
 CC disorders. The MEMAP polynucleotides and proteins are also used for the
 CC diagnosis of these disorders. Specific examples of these disorders
 CC include cancer, inflammation, atherosclerosis, epilepsy and diarrhoea.
 CC MEMAP proteins can be used to screen for compounds which specifically
 CC bind MEMAP including antibodies, oligonucleotides, proteins and small
 CC molecules. MEMAP polynucleotides can be used to prepare transgenic
 CC animals which can be studied to provide information concerning human
 CC disease. Anti-MEMAP antibodies are useful in immunoassays for the
 CC detection of MEMAP protein and can be used as antagonists to treat or
 CC prevent a disorder associated with MEMAP. Polynucleotides encoding MEMAP
 CC can be delivered to target cells with genetic abnormalities with respect
 CC to the expression of MEMAP to treat or prevent a disorder associated with
 CC MEMAP
 XX
 XX
 SQ Sequence 353 AA;

Query Match 38.6%; Score 295.5; DB 4; Length 353;
 Best Local Similarity 46.3%; Pred. NO. 1.5e-23;
 Matches 62; Conservative 29; Mismatches 36; Indels 7; Gaps 2;
 QY 1 DLHIVQADNPISVVEECLEAFKQVFGSLESRRTAQVRYLKPQBEQKVSAYVLRLETL 60
 Db 212 DVIRILKSNPAITTAELKALEQVFGSVSSRDQIKFLNTYQNPGEKLSAYVIRLEPL 271
 QY 61 LRRAVEKRAIPRRIADQVRLEQVMAGA----TLNQMLWCRLRELKQDQPPPSFLELMKVI 116
 Db 272 LQKVEKGADKQNVQARLEQVITAGANHSQAIRQLWL---TGAGGEPAPNLFQILLVQI 328
 QY 117 REEEEPASFENES 130
 Db 329 REBEAKKEEEAE 342

RESULT 8
 AAO16179
 ID AAO16179 standard; protein; 452 AA.
 AC AAO16179;
 DT 28-MAR-2003 (first entry)
 DE Human protein #5.
 XX
 XX Human; vaccine; adult whole brain; foetal whole brain; tonsil;
 KW adult hippocampus; disease-associated SNP analysis; knockout mouse;
 KW disease model mouse; cancer; neurological disorder.
 XX
 OS Homo sapiens.
 XX
 XX WO200299103-A1.
 PN
 PD 12-DEC-2002.
 XX
 XX 27-MAY-2002; 2002WO-JP005134.
 XX
 XX 04-JUN-2001; 2001JP-00168370.
 PR
 XX 16-AUG-2001; 2001JP-00246915.
 XX
 XX (KAZU-) KAZUSA DNA RES INST FOUND.
 PA (PROT-) PROTEIN EXPRESS CO LTD.
 XX
 XX Ohara O, Nagase T, Nakajima D;
 PI
 XX WPI; 2003-140622/13.
 DR N-PSDB; AAL51207.
 XX
 XX DNA preferentially expressed in human adult and fetal brain tissue useful
 PT for diagnosis, treatment and analysis of cancer and mental disorders.
 PT
 XX
 XX Claim 1; Page 56-60; 73pp; Japanese.

CC The invention comprises the amino acid and coding sequences of seven
 CC human proteins that are preferentially expressed in adult whole brain,
 CC foetal whole brain, tonsil and adult hippocampus tissue. The DNA
 CC sequences are useful for the analysis of disease-associated single
 CC nucleotide polymorphisms and the production of knockout and human disease
 CC model mice. The DNA and protein sequences of the invention are useful for
 CC the prevention (vaccine) and treatment of cancer and neurological
 CC disorders. The present amino acid sequence represents a human protein of
 CC the invention
 XX
 XX Sequence 452 AA;

Query Match 37.5%; Score 287; DB 6; Length 452;
 Best Local Similarity 43.8%; Pred. NO. 1.8e-22;
 Matches 63; Conservative 30; Mismatches 43; Indels 8; Gaps 2;
 QY 2 LMHIVQADNPISVVEECLEAFKQVFGSLESRRTAQVRYLKPQBEQKVSAYVLRLETL 61
 Db 215 IMRVQANDSITVEQCLDAKQIFGDKEDFRASQFRPLQTSPIKGEKVSFLLRLEPL 274
 QY 62 RRAVEKRAIPRRIADQVRLEQVMAGATLNQMLWCRLRELKQDQPPPSFLELMKVIRREEE 121
 Db 275 QKAVHKSPLSVSRSTDMIRLKHLLARVAMTPALRGKLELLDQRCPPNPFLELMKLRDEEE 334
 QY 122 ---EEASFENESIEEPERDQYGR 142
 Db 335 WENTEAVMNK-----EKPSGRGR 353

RESULT 9
 ADA54410
 ID ADA54410 standard; protein; 399 AA.
 XX
 AC ADA54410;
 XX
 XX 20-NOV-2003 (first entry)
 DT
 DE Human protein, SEQ ID 1978.
 XX
 XX Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
 KW Gene Therapy; human; secretory protein; membrane proteins; cancer;
 KW inflammatory disease; osteoporosis; neurological disease.
 XX
 OS Homo sapiens.
 XX
 XX EP1293569-A2.
 PN
 PD 19-MAR-2003.
 XX
 XX 21-MAR-2002; 2002EP-00006586.
 PF
 XX 14-SEP-2001; 2001JP-00328381.
 PR
 XX 24-JAN-2002; 2002US-0350435P.
 XX
 XX (HELI-) HELIX RES INST.
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX
 XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
 XX
 XX WPI; 2003-395539/38.
 DR N-PSDB; ADA52771.
 XX
 XX New polynucleotides encoding full-length polypeptides, e.g. secretory
 PT and/or membrane proteins, useful for developing medicines for diseases in
 PT which the gene is involved, or as target molecules for gene therapy.
 XX
 XX Claim 14; SEQ ID NO 1978; 205pp; English.
 PS
 XX The present invention relates to novel human secretory or membrane
 CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
 CC ADA54071). The coding sequences are useful in the gene therapy of

disorders, osteoporosis, osteoarthritis, bone degenerative disorders, periodontal disease, liver fibrosis, infections (e.g. viral, fungal or bacterial) or autoimmune diseases (e.g. diabetes, atopic dermatitis). Sequences ABG9988-ABG9989 and ABU0010-ABU00433 represent human polypeptides of the invention. Note: The sequence data for this patent is not represented in the printed specification but is based on sequence information supplied by the European Patent Office

Query Match 36.2%; Score 277; DB 6; Length 399;
Best Local Similarity 43.5%; Pred. No. 1.9e-21;
Matches 60; Conservative 30; Mismatches 21; Indels 0; Gaps 0;

2 LMHWQADNPISVBECEAFKQVFGSLESRRTAQVRLKPYQEGEKVSAYVLRLETL 61
208 LVHALLAENPARTAQDCALAAQVFGDNESQATIRVKCLTAQQQSGERLSAFVLRLEVL 267

62 BRAVEKRAIPRIADQVRLQVMAQATNQMLWCRLRELKDGPPPPFLEMLKVVREEE 121
268 QKAMEKEALASADRVLRQMLTRAHLTEPLDEALRKLRMAGRSFLEMLGLVRESEA 327

122 BEASFENESIEPBERDG 139
328 WEASLARSVRAQTORGAG 345

RESULT 11
ID AAE01336 standard; protein; 280 AA.
XX AAE01336;
XX 17-JUL-2001 (first entry)
XX Human gene 22 encoded secreted protein fragment, SEQ ID NO:201.
XX Human; secreted protein; proliferative disorder; cancer; tumour;
XX foetal abnormality; developmental abnormality; haematopoietic disorder;
XX immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
XX inflammation; allergy; neurological disorder; Alzheimer's disease;
XX Parkinson's disease; cognitive disorder; schizophrenia; asthma;
XX skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
XX cardiovascular disorder; angiotensin-related disorder;
XX gastrointestinal disorder; pregnancy-related disorder;
XX endocrine disorder; infection; wound healing; vulvular; cell culture;
XX chemotaxis; food additive; gene therapy; binding partner identification.
XX Homo sapiens.
XX OS
XX WO200134769-A2.
XX 17-MAY-2001.
XX 01-NOV-2000; 2000WO-US030040.
XX 05-NOV-1999; 99US-0163580P.
XX 30-JUN-2000; 2000US-0215130P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Ruben SM, Komatsoulis GA, Wei P, Baker KP, Fiscella M;
XX WPI; 2001-308781/32.
XX New isolated nucleic acid molecule encoding a human secreted protein is
XX used in preventing, treating or ameliorating a medical condition.
XX Disclosure; Page 46; 519pp; English.
XX AAD05121-AA005203 represent cDNAs corresponding to 24 human secreted
XX protein genes, and AAE01232-AAE01311 represent the proteins they encode.
XX AAE01312-AAE01340 represent human secreted protein variants or fragments.

diseases caused by abnormalities of the proteins, e.g. cancer, inflammatory diseases, osteoporosis or neurological disease.

Query Match 36.2%; Score 277; DB 6; Length 399;
Best Local Similarity 43.5%; Pred. No. 1.9e-21;
Matches 60; Conservative 30; Mismatches 48; Indels 0; Gaps 0;

2 LMHWQADNPISVBECEAFKQVFGSLESRRTAQVRLKPYQEGEKVSAYVLRLETL 61
208 LVHALLAENPARTAQDCALAAQVFGDNESQATIRVKCLTAQQQSGERLSAFVLRLEVL 267

62 BRAVEKRAIPRIADQVRLQVMAQATNQMLWCRLRELKDGPPPPFLEMLKVVREEE 121
268 QKAMEKEALASADRVLRQMLTRAHLTEPLDEALRKLRMAGRSFLEMLGLVRESEA 327

122 BEASFENESIEPBERDG 139
328 WEASLARSVRAQTORGAG 345

RESULT 10
ID ABG99947 standard; protein; 399 AA.
XX ABG99947;
XX 17-JAN-2003 (first entry)
XX Human novel polypeptide #60.
XX Human; genetic disorder; gene mapping; medical imaging; cancer;
XX neurodegenerative disorder; lymphoid cell disorder; osteoporosis;
XX Parkinson's disease; Alzheimer's disease; bone degenerative disorder;
XX osteoarthritis; periodontal disease; liver fibrosis; viral infection;
XX fungal infection; bacterial infection; autoimmune disease; diabetes;
XX atopic dermatitis.
XX Homo sapiens.
XX OS
XX WO200274961-A1.
XX 26-SEP-2002.
XX 14-MAR-2002; 2002WO-US005109.
XX 15-MAR-2001; 2001US-00810173.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Zhou P, Goodrich R, Asundi V, Zhang J, Zhao QA, Ren F;
XX Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
XX Wehrman T, Wang J, Wang D, Drmanac RT;
XX WPI; 2003-040556/03.
XX N-PSDB; ABX05045.
XX New isolated polypeptides and polynucleotides, useful for preventing,
XX treating or ameliorating medical conditions, such as cancer,
XX neurodegenerative disorders, lymphoid cell disorders, bone degenerative
XX disorders, and infections.
XX Claim 9; SEQ ID NO 586; 235pp; English.
XX The invention relates to human polynucleotides and the polypeptides they
XX encode. The polynucleotides and polypeptides are useful in diagnostics,
XX forensics, gene mapping, medical imaging, identification of mutations
XX responsible for genetic disorders or other traits, assessing biodiversity
XX and producing many other types of data and products dependent on DNA and
XX amino acid sequences. They are also useful for preventing, treating or
XX ameliorating medical conditions, such as cancer, neurodegenerative
XX disorders (e.g. Parkinson's disease, Alzheimer's disease), lymphoid cell

CC The secreted proteins and their genes are useful for preventing, treating
 CC or ameliorating medical conditions, e.g., by protein or gene therapy.
 CC Pathological conditions can be diagnosed by determining the amount of the
 CC new protein in a sample or by determining the presence of mutations in
 CC the new genes. Specific uses are described for each of the 24 genes,
 CC based on the tissues in which they are most highly expressed, and include
 CC developing products for the diagnosis or treatment of proliferative
 CC disorders, cancer, tumours, foetal and developmental abnormalities,
 CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
 CC diseases (e.g., rheumatoid arthritis), inflammation, allergies,
 CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
 CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
 CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
 CC angiogenic disorders, kidney disorders, gastrointestinal disorders,
 CC pregnancy-related disorders, endocrine disorders, and infections. The
 CC proteins can also be used to aid wound healing and epithelial cell
 CC proliferation, to prevent skin aging due to sunburn, to maintain organs
 CC before transplantation, for supporting cell culture of primary tissues,
 CC to regenerate tissues, to identify their cognate ligands or binding
 CC partners, and in chemotaxis, and can be used as a food additive or
 CC preservative to modify storage properties. Antibodies specific for a
 CC protein of the invention can be used in alleviating symptoms associated
 CC with the disorders mentioned above, and in diagnostic immunoassays e.g.,
 CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The
 CC present sequence represents a human secreted protein fragment referred to
 CC in the disclosure of the invention
 CC
 CC SQ Sequence 280 AA;

Query Match 35.4%; Score 271.5; DB 4; Length 280;
 Best Local Similarity 44.6%; Pred. No. 4.7e-21;
 Matches 58; Conservative 32; Mismatches 29; Indels 11; Gaps 3;

QY 1 DLMHIVQADNPSTVBECELEAFQVFGSLESRTAQVRYLKPQOEGEKVSAYVLRLETL 60
 Db 145 DVIRVLKINNPLITVDECLQALEEVFGVTDNPRELQVKYLTYYQKDEKLSAYVLRLEPL 204
 QY 61 LRRAVEKRAIPRIADQVRLEQVMAGA---TLNQMLWCLRELKDGQPPSPFLELMKVIR 117
 Db 205 LQKLVQGAERDAVNAQRLDQVIAGAVHKTIRREL-----NLPEDGAPGFLQLLVLIK 259
 QY 118 E---EEEEEA 124
 Db 260 DYEAEEEEEA 269

RESULT 12

AAE01340
 ID AAE01340 standard; protein; 351 AA.

XX AAE01340;

XX 17-JUL-2001 (first entry)

DE Human gene 22 encoded secreted protein fragment, SEQ ID NO:205.

KW Human; secreted protein; proliferative disorder; cancer; tumour;
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KW inflammation; allergy; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
 KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
 KW cardiovascular disorder; angiogenic disorder; kidney disorder;
 KW gastrointestinal disorder; pregnancy-related disorder;
 KW endocrine disorder; infection; wound healing; vulnerability; cell culture;
 KW chemotaxis; food additive; gene therapy; binding partner identification.

OS Homo sapiens.

XX

PN WO200134769-A2.

XX

PD 17-MAY-2001.

XX

PF 01-NOV-2000; 2000MO-US030040.

XX

PR 05-NOV-1999; 99US-0163580P.

PR 30-JUN-2000; 2000US-0215130P.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX

PI Ruben SM, Komatsoulis GA, Wei P, Baker KP, Fiscella M;

XX WPI; 2001-308781/32.

XX

PT New isolated nucleic acid molecule encoding a human secreted protein is

XX used in preventing, treating or ameliorating a medical condition.

XX

XX Disclosure; Page 46; 519pp; English.

XX

CC AAD05121-AAD05203 represent cDNAs corresponding to 24 human secreted

CC protein genes, and AAE01232-AAE01311 represent the proteins they encode.

CC AAE01312-AAE01340 represent human secreted protein variants or fragments.

CC The secreted proteins and their genes are useful for preventing, treating

CC or ameliorating medical conditions, e.g., by protein or gene therapy.

CC Pathological conditions can be diagnosed by determining the amount of the

CC new protein in a sample or by determining the presence of mutations in

CC the new genes. Specific uses are described for each of the 24 genes,

CC based on the tissues in which they are most highly expressed, and include

CC developing products for the diagnosis or treatment of proliferative

CC disorders, cancer, tumours, foetal and developmental abnormalities,

CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune

CC diseases (e.g., rheumatoid arthritis), inflammation, allergies,

CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),

CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,

CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,

CC angiogenic disorders, kidney disorders, gastrointestinal disorders,

CC pregnancy-related disorders, endocrine disorders, and infections. The

CC proteins can also be used to aid wound healing and epithelial cell

CC proliferation, to prevent skin aging due to sunburn, to maintain organs

CC before transplantation, for supporting cell culture of primary tissues,

CC to regenerate tissues, to identify their cognate ligands or binding

CC partners, and in chemotaxis, and can be used as a food additive or

CC preservative to modify storage properties. Antibodies specific for a

CC protein of the invention can be used in alleviating symptoms associated

CC with the disorders mentioned above, and in diagnostic immunoassays e.g.,

CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The

CC present sequence represents a human secreted protein fragment referred to

CC in the disclosure of the invention
 CC
 CC SQ Sequence 351 AA;

Query Match 35.4%; Score 271.5; DB 4; Length 351;

Best Local Similarity 44.6%; Pred. No. 6.4e-21;

Matches 58; Conservative 32; Mismatches 29; Indels 11; Gaps 3;

QY 1 DLMHIVQADNPSTVBECELEAFQVFGSLESRTAQVRYLKPQOEGEKVSAYVLRLETL 60

Db 216 DVIRVLKINNPLITVDECLQALEEVFGVTDNPRELQVKYLTYYQKDEKLSAYVLRLEPL 275

QY 61 LRRAVEKRAIPRIADQVRLEQVMAGA---TLNQMLWCLRELKDGQPPSPFLELMKVIR 117

Db 276 LQKLVQGAERDAVNAQRLDQVIAGAVHKTIRREL-----NLPEDGAPGFLQLLVLIK 330

QY 118 E---EEEEEA 124

Db 331 DYEAEEEEEA 340

RESULT 13

AAE01340

ID AAE01340 standard; protein; 351 AA.

XX AAE01340;

XX 12-JUN-2001 (first entry)

XX

DE Human membrane associated protein MEMAP-1.
XX Human; membrane associated protein; MEMAP; diagnosis; cytostatic;
KW antiinflammatory; anticonvulsant; immunosuppressive; antiarrhythmic;
KW antiarteriosclerotic; gene therapy; cell proliferative disorder;
KW autoimmune disorder; inflammatory disorder; neurological disorder;
KW gastrointestinal disorder; cancer; inflammation; atherosclerosis;
KW epilepsy; diarrhoea.
XX Homo sapiens.
XX WO200112662-A2.
XX 22-FEB-2001.
XX 14-AUG-2000; 2000WO-US022315.
XX 17-AUG-1999; 99US-0149641P.
XX 09-NOV-1999; 99US-0164203P.
XX (INCY-) INCYTE GENOMICS INC.
XX Lal P, Yue H, Tang YT, Bandman O, Burford N, Azimzai Y;
PI Baughn MR, Lu DAM, Patterson C;
XX WPI; 2001-168860/17.
XX N-PSDB; AAF81741.
XX Isolated polypeptide with a human membrane associated protein sequence is
PT useful for the diagnosis, prevention and treatment of cell proliferative,
PT autoimmune/inflammatory, neurological and gastrointestinal disorders.
XX Claim 1; Page 114-115; 173pp; English.
XX AAF81741 to AAF81777 encode the human membrane associated proteins
CC (MEMAP) given in AAB74695 to AAB74731. MEMAPs have cytostatic,
CC antiinflammatory, anticonvulsant, immunosuppressive, antiarrhythmic and
CC antiarteriosclerotic activities, which can be used in gene therapy.
CC MEMAPs and agonist of MEMAPs can be used to treat a disease or condition
CC associated with decreased expression of functional MEMAP and antagonists
CC of MEMAP are used to treat a disease or condition associated with
CC overexpression of functional MEMAP. These disorders include cell
CC proliferative, autoimmune/inflammatory, neurological and gastrointestinal
CC disorders. The MEMAP polynucleotides and proteins are also used for the
CC diagnosis of these disorders. Specific examples of these disorders
CC include cancer, inflammation, atherosclerosis, epilepsy and diarrhoea.
CC MEMAP proteins can be used to screen for compounds which specifically
CC bind MEMAP including antibodies, oligonucleotides, proteins and small
CC molecules. MEMAP polynucleotides can be used to prepare transgenic
CC animals which can be studied to provide information concerning human
CC disease. Anti-MEMAP antibodies are useful in immunoassays for the
CC detection of MEMAP protein and can be used as antagonists to treat or
CC prevent a disorder associated with MEMAP. Polynucleotides encoding MEMAP
CC can be delivered to target cells with genetic abnormalities with respect
CC to the expression of MEMAP to treat or prevent a disorder associated with
CC MEMAP
XX
SQ

Query Match 35.4%; Score 271.5; DB 4; Length 351;
Best Local Similarity 44.6%; Pred. No. 6.4e-21;
Matches 58; Conservative 33; Mismatches 29; Indels 11; Gaps 3;
QY 1 DLMHIVQADNPSTVSECLAFQVKGSLRSTQVRVLYKPQEGEKVSAYVLETL 60
DB 216 DVIRVLKINPLITVDECLQALBEVFGVTDNPRELQVKYLTYYKDBEKL SAYVLEPL 275
QY 61 LRRAVEKRAIPRIADVRLEQVMAGA---TLNQMVLWRLRELKDQPPSPFLELMKVIR 117
DB 276 LQKLVRGALIERDAVNGARLDQVITAGVHTITREL-----NLPEGPAGFFLQLLVLIK 330
QY 118 E---EEEEEA 124
; |||||

Db 331 DYEAEEEE 340

RESULT 14
AAU08664
ID AAU08664 standard; protein; 351 AA.
XX AC AAU08664;
XX 18-DEC-2001 (first entry)
XX Human NOV8 protein.
XX Human; NOV8; cytostatic; nootropic; neuroprotective; vulnary;
KW cerebroprotective; antiparkinsonian; hypotensive; antiasthmatic;
KW antidiabetic; antipsoriatic; antiinflammatory; immunosuppressive;
KW antiatherosclerotic; dermatological; cancer; neurological disorder;
KW Parkinson's disease; diabetes mellitus; asthma; enamel defect;
KW immune disorder; autoimmune disease; respiratory disorder; bone disorder;
KW musculoskeletal disorders; leukaemia; lymphoma;
KW cell growth regulation disorder; lesional psoriatic skin;
KW atherosclerosis; abdominal aortic aneurysm.
XX Homo sapiens.
XX WO200168851-A2.
XX 20-SEP-2001.
XX 12-MAR-2001; 2001WO-US007735.
XX 10-MAR-2000; 2000US-0188277P.
XX 10-MAR-2000; 2000US-0188316P.
XX 14-MAR-2000; 2000US-0189139P.
XX 14-MAR-2000; 2000US-0189140P.
XX 17-MAR-2000; 2000US-0190231P.
XX 17-MAR-2000; 2000US-0190401P.
XX (CURA-) CURAGEN CORP.
XX Padigar M, Vernet CAM, Fernandes E, Shimkets RA, Spaderna SK;
PI Majumder K, Li L;
XX WPI; 2001-570869/64.
XX N-PSDB; AAS13342.
XX Novel polypeptides and nucleic acids homologous to members of collagen,
PT potassium channel, tuftelin family of proteins for diagnosing, treating
PT cancer, atherosclerosis, neurological, skin and enamel defect disorders.
XX Claim 1; Page 29; 128pp; English.
XX The invention relates to isolated NOVX (NOVX1-11) polypeptides and the
CC polynucleotides that encode them. NOVX polypeptides, polynucleotides and
CC anti-NOVX antibodies are useful for treating or preventing a pathology
CC associated with NOVX polypeptide in humans and for treating a syndrome
CC associated with human disease e.g. disorders characterised by altered
CC cell motility, proliferation and migration e.g. cancer, angiogenesis and
CC wound healing (NOV1-3), neurological disorders, e.g. episodic ataxia,
CC autosomal dominant myokymia, stroke, Parkinson's disease, Alzheimer's
CC disease, non-insulin dependent diabetes mellitus, asthma, hypertension
CC and seizure (NOV4), enamel defects, such as amelogenesis imperfecta and
CC disorders involving enamel defects, including hypoplasia and
CC hypomineralisation (NOV5-7), paraneoplastic neurological disorders, e.g.
CC paraneoplastic limbic of brain-stem encephalitis occurring during
CC testicular cancer, diabetes, reproductive health, metabolic and endocrine
CC disorders, gastrointestinal disorders, immune disorders and autoimmune
CC diseases, respiratory disorders, bone disorders, musculoskeletal
CC disorders, leukaemia/lymphoma and tissue/cell growth regulation disorders
CC (NOV8), lesional psoriatic skin (NOV9-10) and atherosclerosis, abdominal
CC aortic aneurysm and neurological disorders (NOV11). NOVX polypeptide is
CC also useful for identifying an agent that binds to it and a cell
CC expressing NOVX polypeptide is useful for identifying a therapeutic agent

CC for use in treatment of a NOVX related pathology. The antibodies and a
 CC polypeptide having 95% sequence identity to NOVX polypeptide are useful
 CC for treating a pathological state in a mammal. The present sequence
 CC represents NOV8, a possible neuronal antigen-like protein

XX SQ Sequence 351 AA;

Query Match 35.4%; Score 271.5; DB 4; Length 351;
 Best Local Similarity 44.6%; Pred. No. 6.4e-21;
 Matches 58; Conservative 32; Mismatches 29; Indels 11; Gaps 3;
 QY 1 DLMHIVQADNPISIVVECLAFKQVFGSLERRTAQVRLKPYOEGEKVSAYVRLLETL 60
 Db 216 DVIRVLKINPLITVDECLQALEEVFGVTDNPRLQVKYLTYYQDEKLSAYVRLLEPL 275
 QY 61 LRRRAVEKRAIPRRITADQVRLEQVWAGA---TLNQLMCLRELKQDQPPSPFLEIMKVIR 117
 Db 276 LQKLVQGAERDAVNQARLDQVIAGAVHKTIRREL-----NLPEDGAPGFLQLLVLIK 330
 QY 118 E---EEEEEA 124
 Db 331 DYEAEEEEEA 340

RESULT 15

AAB94854
 ID AAB94854 standard; protein; 237 AA.

AC AAB94854;

XX 26-JUN-2001 (first entry)

XX Human protein sequence SEQ ID NO:16040.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-00116126.

XX 29-JUL-1999; 99JP-00248036.

XX 27-AUG-1999; 99JP-00300253.

XX 11-JAN-2000; 2000JP-00118776.

XX 02-MAY-2000; 2000JP-00183767.

XX 09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

XX Claim 8; SEQ ID NO 16040; 2537pp + Sequence Listing; English.

XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dr primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a

CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention

XX SQ Sequence 237 AA;

Query Match 34.8%; Score 266.5; DB 4; Length 237;
 Best Local Similarity 43.8%; Pred. No. 1.3e-20;
 Matches 57; Conservative 33; Mismatches 29; Indels 11; Gaps 3;

QY 1 DLMHIVQADNPISIVVECLAFKQVFGSLERRTAQVRLKPYOEGEKVSAYVRLLETL 60
 Db 102 DVIRVLKINPLITVDECLQALEEVFGVADNPRLQVKYLTTHQDEKLSAYVRLLEPL 161
 QY 61 LRRRAVEKRAIPRRITADQVRLEQVWAGA---TLNQLMCLRELKQDQPPSPFLEIMKVIR 117
 Db 162 LQKLVQGAERDAVNQARLDQVIAGAVHKTIRREL-----NLPEDGAPGFLQLLVLIK 216
 QY 118 E---EEEEEA 124
 Db 217 DYEAEEEEEA 226

Search completed: September 27, 2004, 17:08:22
 Job time : 29.0768 secs

Blank